Assimilation of Soil Moisture and Land Surface Temperature with the Ensemble Kalman Filter

*Workshop on **Predictability, Observations,** and **Uncertainties** in Geosciences*
Florida State University
March 13-15, 2006

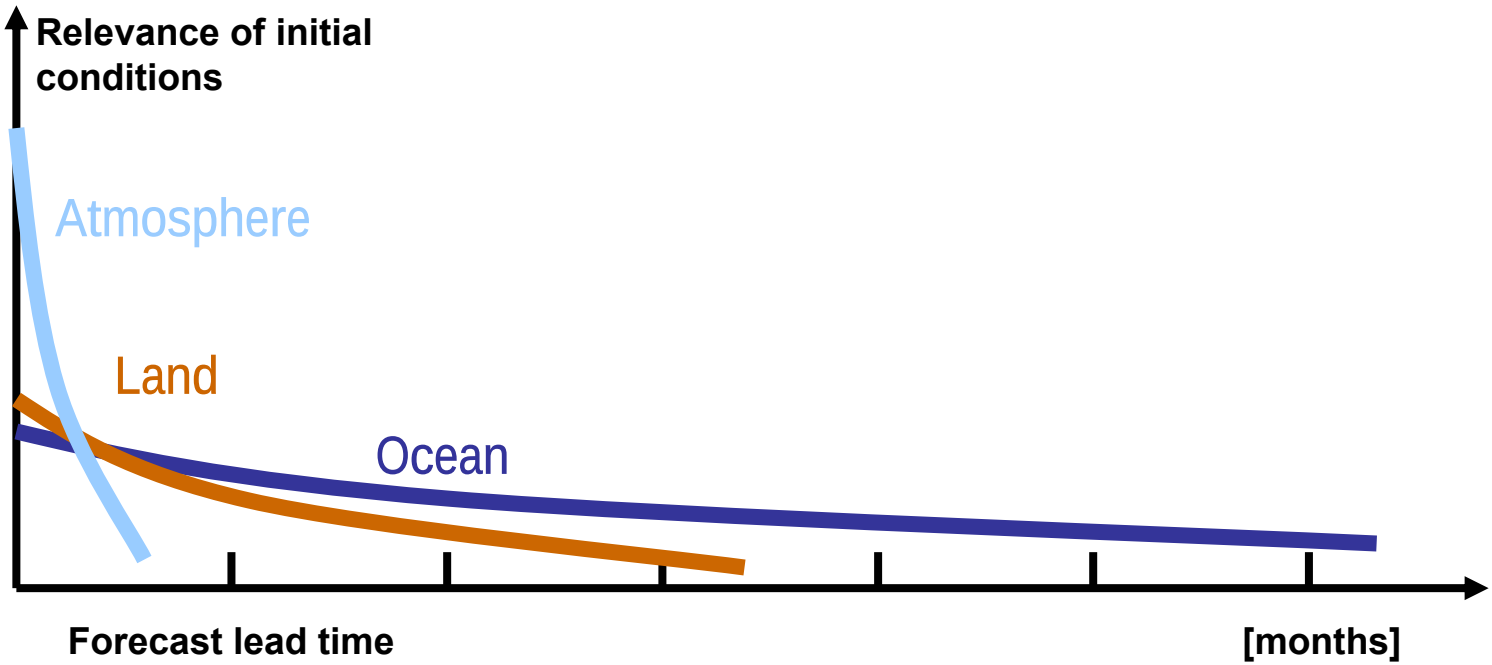
Rolf Reichle[1,2]
Mike Bosilovich[1]
Randy Koster[1]
Ping Liu[1,3]
Sarith Mahanama[1,2]
Jon Radokovich[1,3]

1 – Global Modeling and Assimilation Office, NASA
2 – GEST, University of Maryland, Baltimore County
3 – SAIC

Outline

Motivation	Seasonal climate *predictability* associated with land surface conditions
Method	Land data assimilation with the Ensemble Kalman filter
Soil moisture	Satellite *observations*, *uncertainties* & data assimilation
Land surface temperature	Satellite *observations*, *uncertainties* & data assimilation
Conclusions	

Initial conditions and forecast lead time



Weather prediction limited by atmospheric chaos beyond two weeks.

Seasonal climate prediction must rely on slower moving components of the Earth's system (ocean heat content, soil moisture, snow).

Land initialization relevant at 2 week – 2 month (sub-seasonal) time scale.

A simple view of land-atmosphere feedback

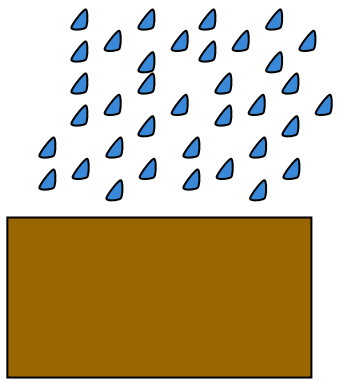
Precipitation wets the surface...

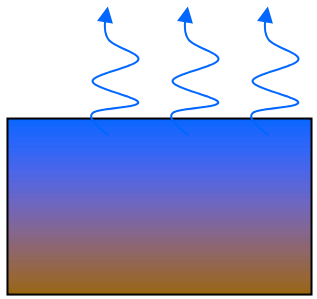
…causing soil moisture to increase...

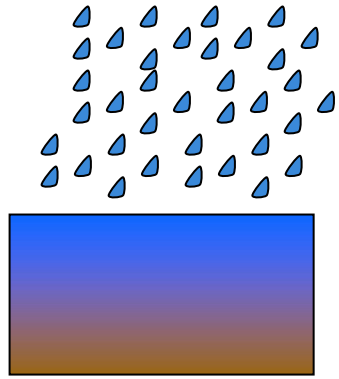
…which causes evaporation to increase during subsequent days and weeks...

…which affects the overlying atmosphere (the boundary layer structure, humidity, etc.)...

…thereby (maybe) inducing additional precipitation







Perhaps such feedback contributes to predictability?

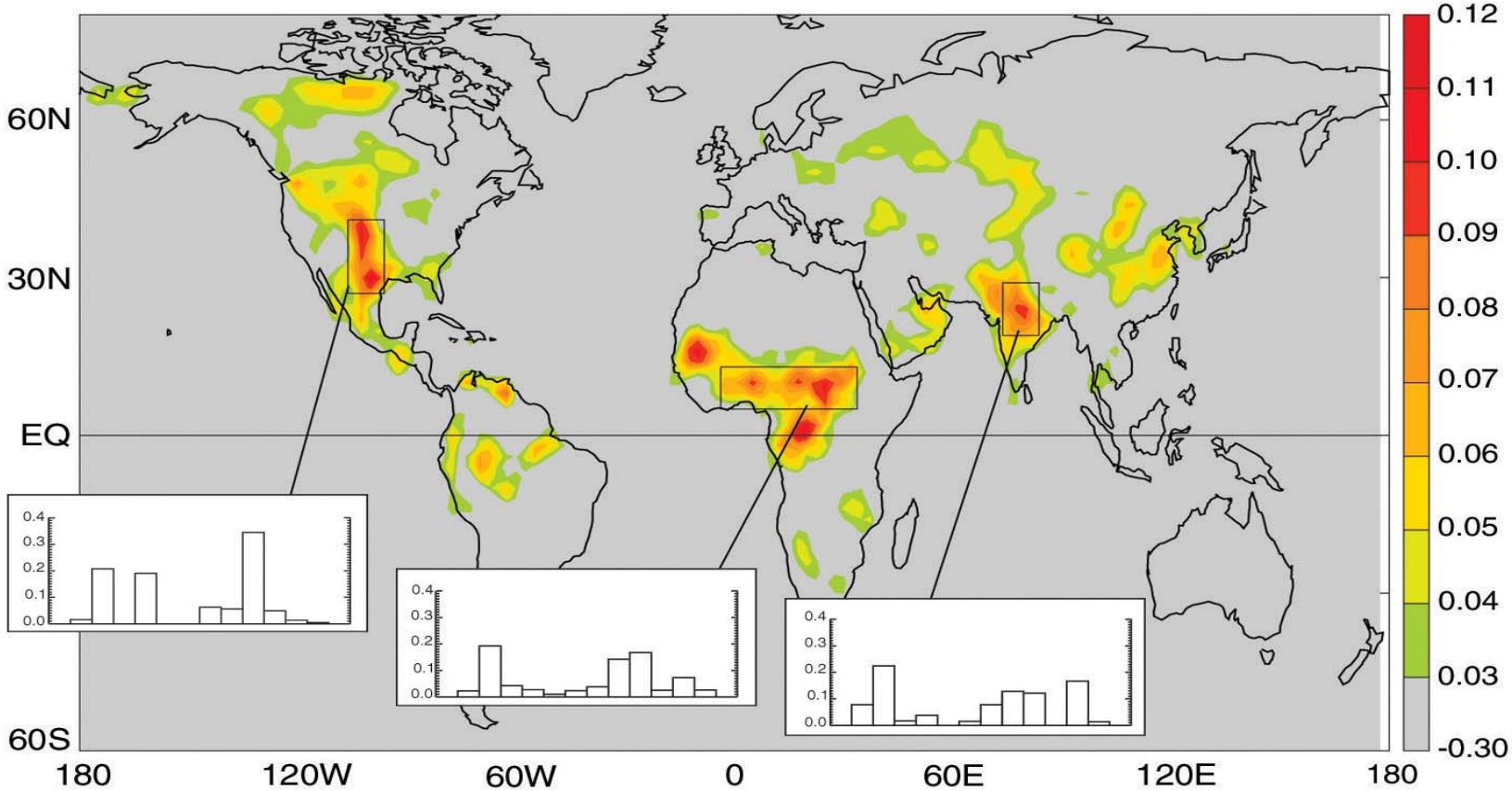
Two things must happen:

1. A soil moisture anomaly must be “remembered” into the forecast period.
2. The atmosphere must respond predictably to soil moisture anomalies.

In other words, need strong land-atmosphere coupling…

Soil moisture memory and "hot spots"

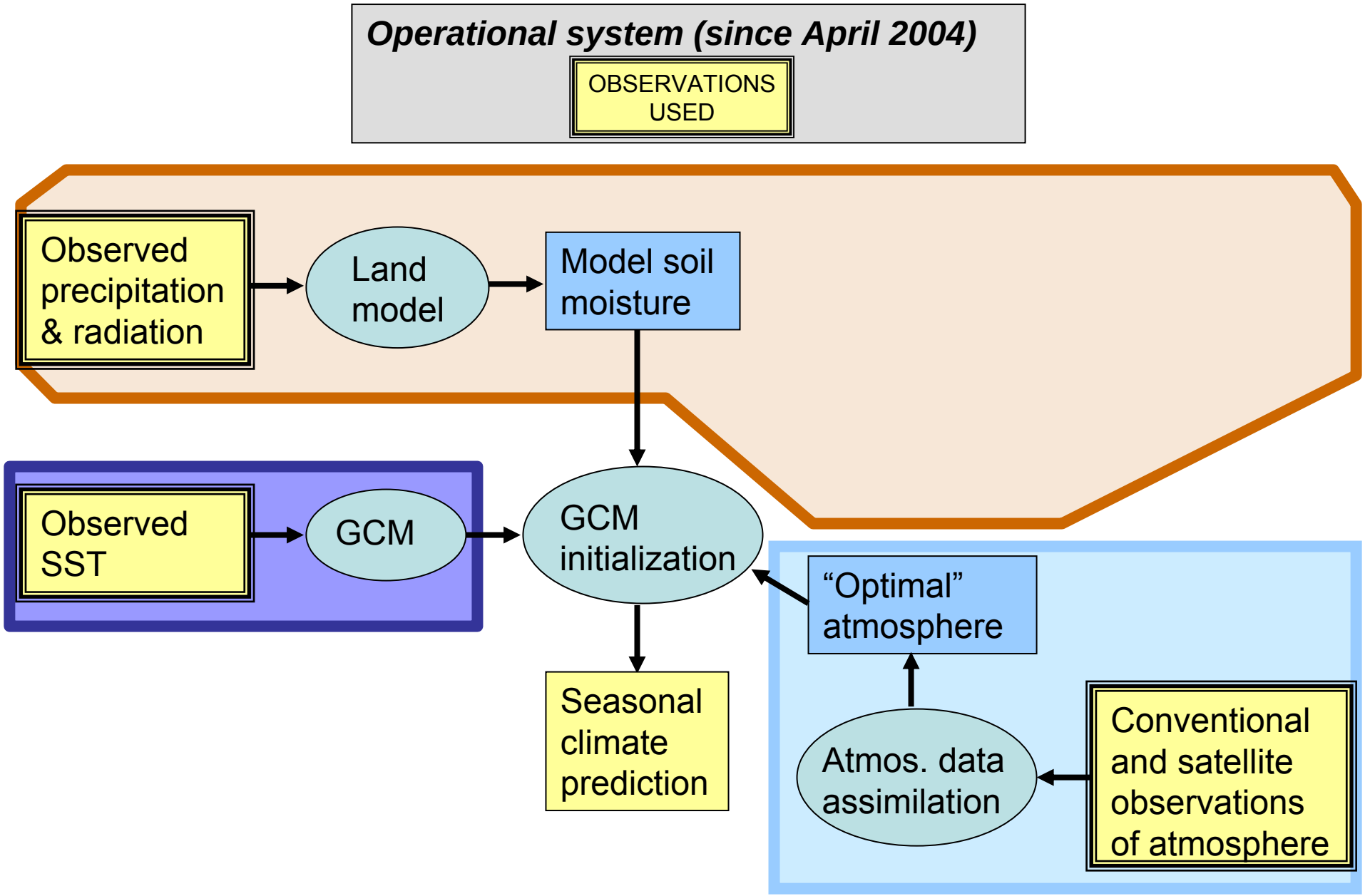
Land-atmosphere coupling strength (JJA), averaged across AGCMs



Koster et. al, *Science*, 2004

"Hot spots" where soil moisture conditions affect summer rainfall (multi-model consensus).

NASA seasonal forecast initialization



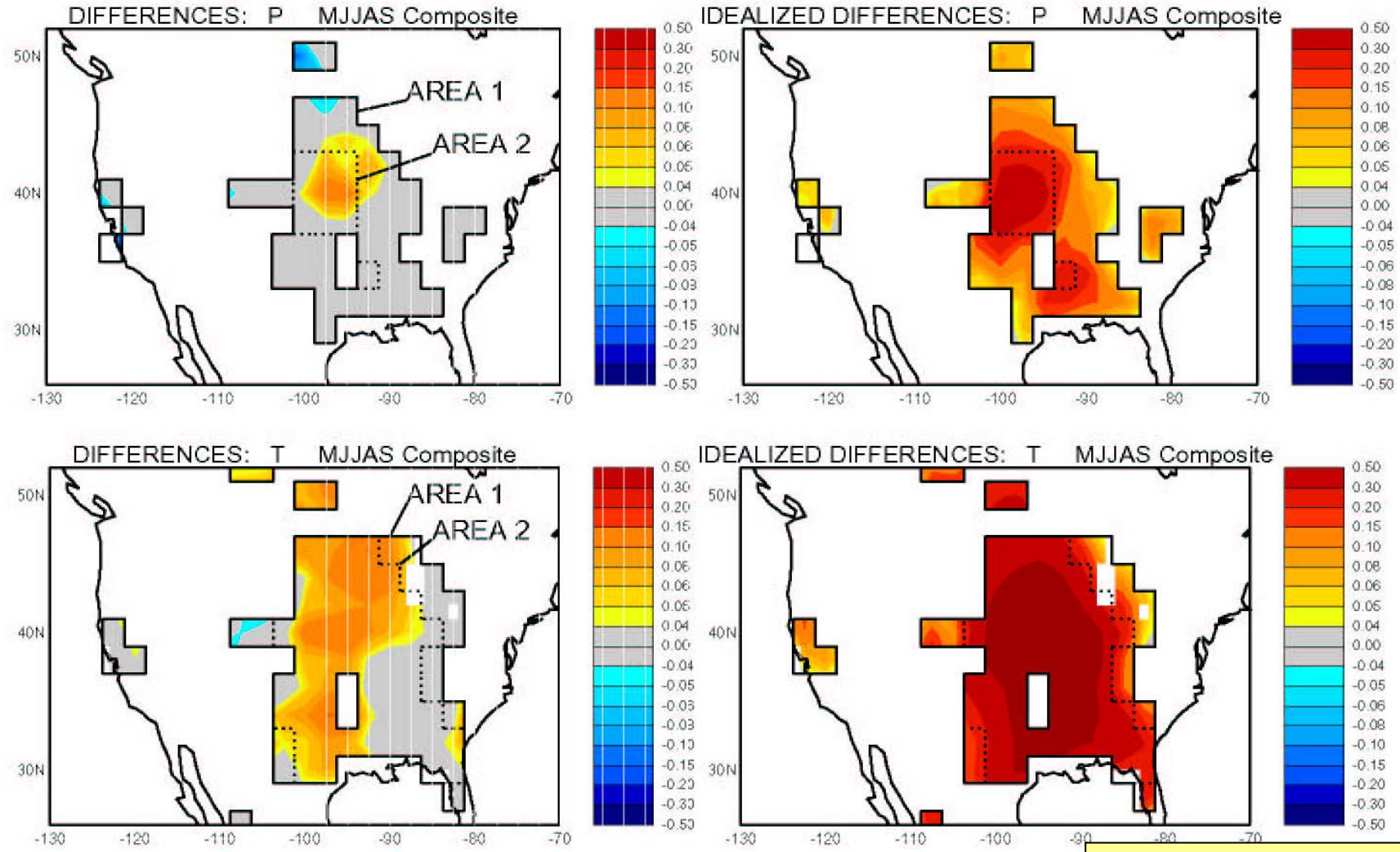
Contribution of land moisture initialization to skill of one-month forecasts

Actual contribution to skill

Maximum possible contribution
(Idealized skill w/ perfect init and model)

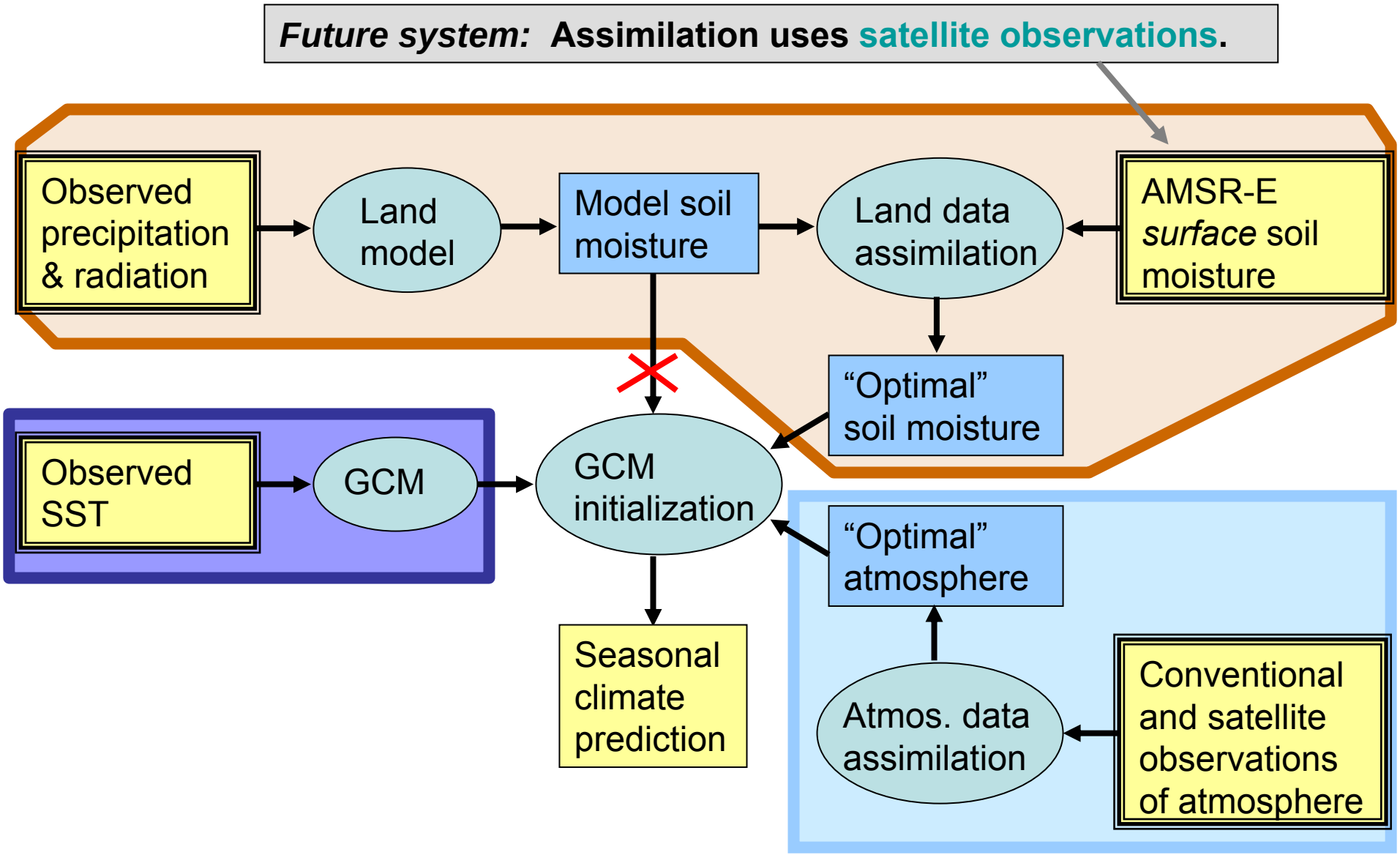
Precipitation

Temperature



Koster et al., JHM, 2004

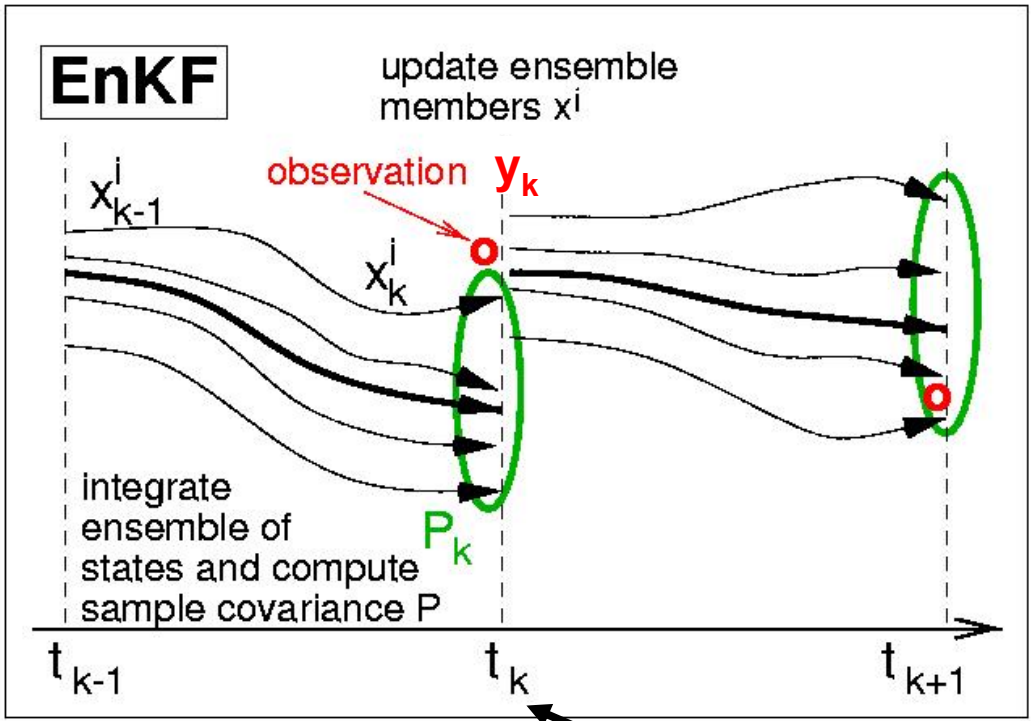
NASA seasonal forecast initialization



Outline

Motivation	Seasonal climate *predictability* associated with land surface conditions
Method	Land data assimilation with the Ensemble Kalman filter
Soil moisture	Satellite *observations*, *uncertainties* & data assimilation
Land surface temperature	Satellite *observations*, *uncertainties* & data assimilation
Conclusions	

Ensemble Kalman filter (EnKF)



Nonlinearly propagates ensemble of model trajectories.

Can account for wide range of model errors (incl. non-additive).

Approx.: **Ensemble size.**

Linearized update.

$\mathbf{x}_k^i$	state vector (eg soil moisture)
$\mathbf{P}_k$	state error covariance
$\mathbf{R}_k$	observation error covariance

Propagation $\mathbf{t}_{k-1}$ to $\mathbf{t}_k$:

$$\mathbf{x}_k^{i+} = \mathbf{f}(\mathbf{x}_{k-1}^{i-}) + \mathbf{w}_k^i$$

$\mathbf{w}$ = model error

Update at $\mathbf{t}_k$:

$$\mathbf{x}_k^{i+} = \mathbf{x}_k^{i-} + \mathbf{K}_k(\mathbf{y}_k^i - \mathbf{x}_k^{i-})$$

for each ensemble member i=1…N

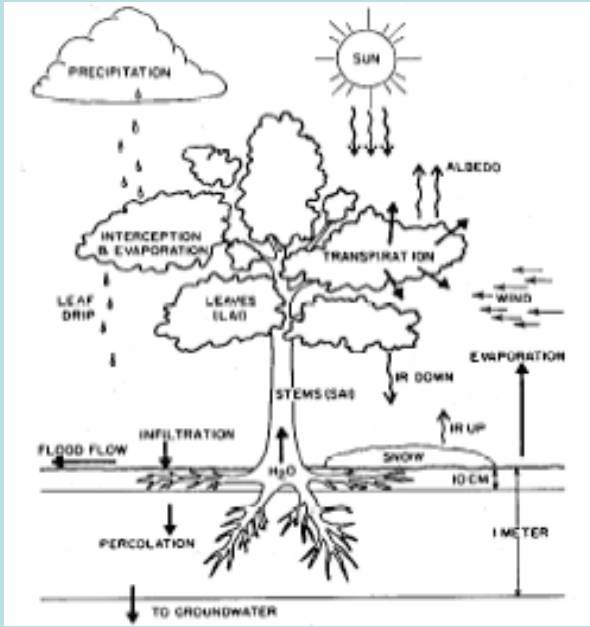
$$\mathbf{K}_k = \mathbf{P}_k (\mathbf{P}_k + \mathbf{R}_k)^{-1}$$

with $\mathbf{P}_k$ computed from ensemble spread

Error structure in land data assimilation

The land model describes the **water and energy balance** at the land surface.

Trajectories of the (uncoupled) land model are determined by surface meteorological forcings, not initial conditions – **no chaos/butterfly effect**.



© Hartmann, Global Physical Climatology

Dominant errors are "model errors" in

(a) surface meteorological forcings (time-varying b.c.'s) and
(b) model parameterizations.

⇒ **Maintain ensemble spread through random fields that are**

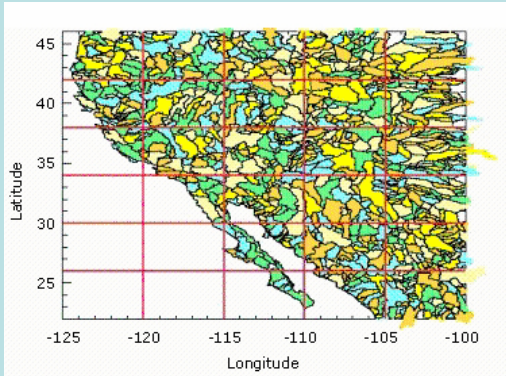
(a) correlated in time (AR1),
(b) correlated in space (Gaussian), and
(c) cross-correlated.

Error structure in land data assimilation

The (uncoupled) land model is a collection of identical “1d” models that are applied to each land surface element (or catchment) independently.

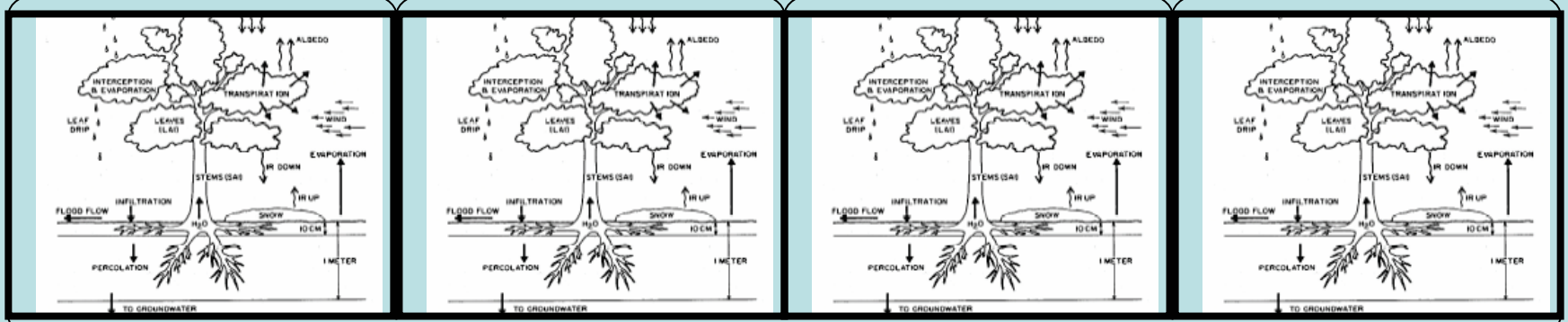
Forcing/model errors are horizontally correlated.

EnKF can be applied
(a) independently to each catchment (“1d-EnKF”) or
(b) simultaneously to a set of catchments (“3d-EnKF”).



“1d-EnKF”

EnKF **EnKF** **EnKF** **EnKF**



“3d-EnKF”

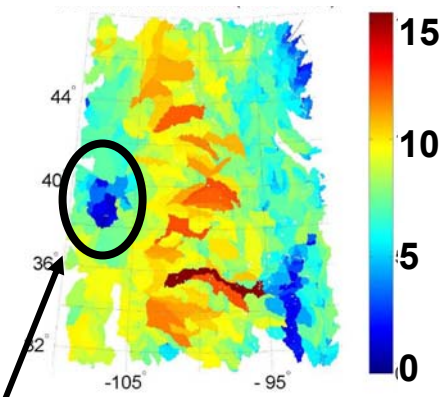
EnKF

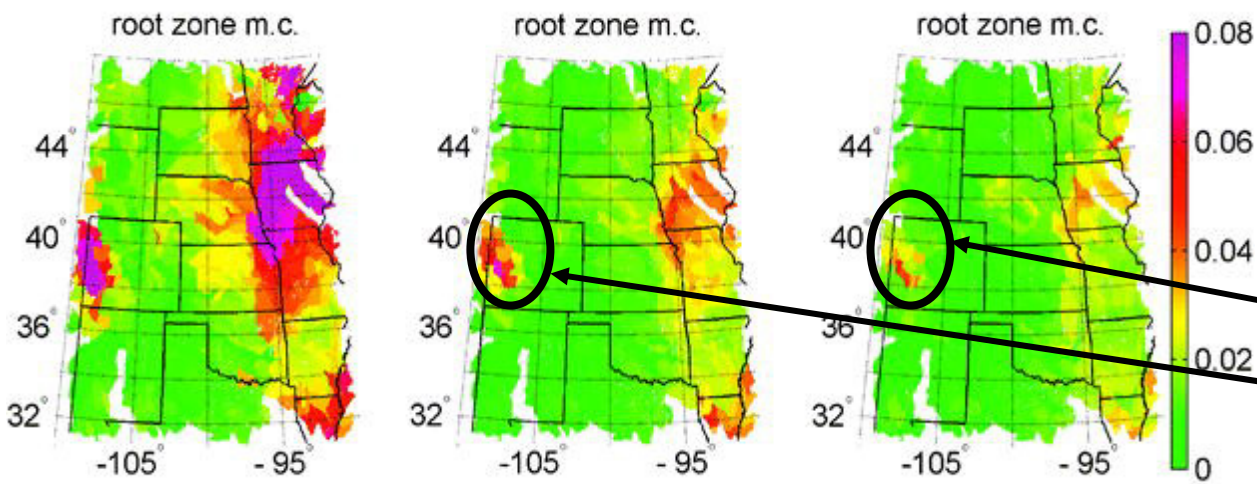
1d-EnKF v. 3d-EnKF

Satellite soil moisture is not available everywhere and always. Can data assimilation spread information horizontally?

Taking horizontal error correlations into account in the 3D-EnKF improves soil moisture estimates over the local (or 1D) EnKF.

 # SMMR data/month





Prior (no assim.) **1D EnKF** **3D EnKF**

Root zone soil moisture estimation errors

3D-EnKF is better than 1D-EnKF where satellite data are sparse.

Reichle & Koster, JHM, 2003

Outline

Motivation	Seasonal climate *predictability* associated with land surface conditions
Method	Land data assimilation with the Ensemble Kalman filter
Soil moisture	Satellite *observations*, *uncertainties* & data assimilation
Land surface temperature	Satellite *observations*, *uncertainties* & data assimilation
Conclusions	

Soil moisture data sources

		"SMMR period" *1979-87 (~8.5 years)*	*"AMSR-E period"* *2002-05 (~3.5 years)*
Soil moisture retrievals	*Sensor*	SMMR (Nimbus 7)	AMSR-E (Aqua)
	Frequency	C-Band (6.6 GHz)	X-Band (10.7 GHz)
	Sampling depth	~1.25 cm	~1 cm
	Horiz. Resolution	~150 km	~40 km
	Equator crossing	12 am/pm	1:30 am/pm
	Algorithm	Owe et al., 2001	Njoku et al. (http://nsidc.org)
Land surface model		NASA Catchment (~0.5°)	(same w/ minor updates)
Meteorol. forcing data (obs.-based)	*Author*	Berg et al., 2005	GLDAS
	Baseline	Re-analysis (ERA-15)	NASA GEOS NWP analysis
	Observations	Monthly	Daily/pentad
	Precipitation	GPCP satellite/gauge	CMAP (5-day)
	Radiation	SRB (1983-87 only)	AGRMET daily
	Air temp./humid.	CRU	(None)
	Horiz. resolution	~2 deg	~2 deg

Soil moisture data sources

		"SMMR period" 1979-87 (~8.5 years)	"AMSR-E period" 2002-05 (~3.5 years)
Soil moisture retrievals	*Sensor*	SMMR (Nimbus 7)	AMSR-E (Aqua)
	Frequency	C-Band (6.6 G	X-Band (10.7
	Sampling depth	~1	
	Horiz. Resolution	5 km	5 km
	Equator crossing	12 am/pm	1:30 am/pm
	Algorithm	Owe et al., 2001	Njoku et al. (http://nsidc.org)
Land surface model		NASA Catchment (~0.5°)	(same w/ minor updates)
Meteorol. forcing data (obs.-based)	*Author*	Berg et al., 2	GLDAS
	Baseline	Re-analy A-15)	NASA GEOS analysis
	Observations	nthly	entad
	Precipitation	CP satellite/gauge	MAP (5-day)
	Radiation	SRB (1983-87 only)	AGRMET daily
	Air temp./humid.	CRU	(None)
	Horiz. resolution	~2 deg	~2 deg

1. SMMR retrievals

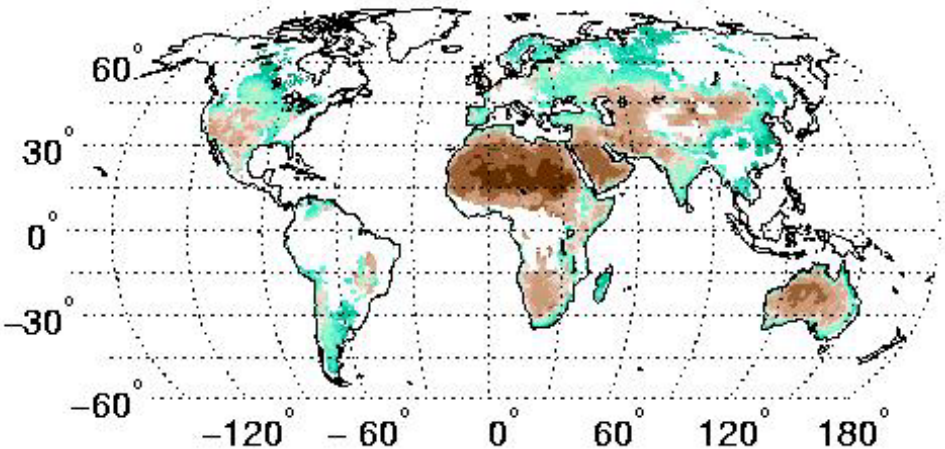
2. AMSR-E retrievals

3. Model (1979-87)

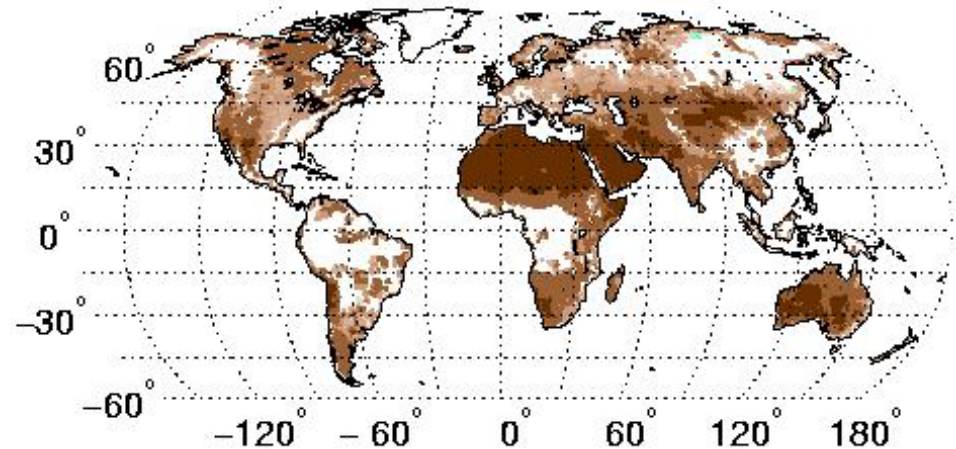
4. Model (2002-05)

Satellite vs. satellite bias (time avg. soil moisture)

SMMR (1979-87)

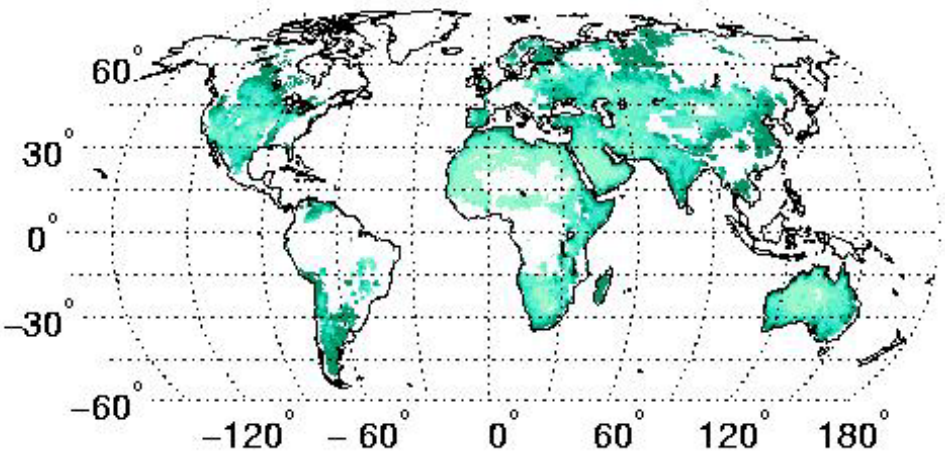


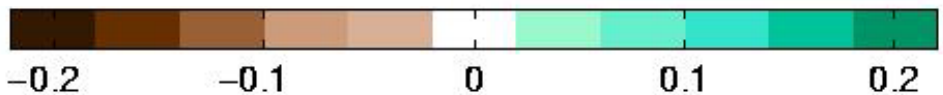
AMSR-E (2002-05)



Soil moisture [m³/m³]

SMMR minus AMSR-E





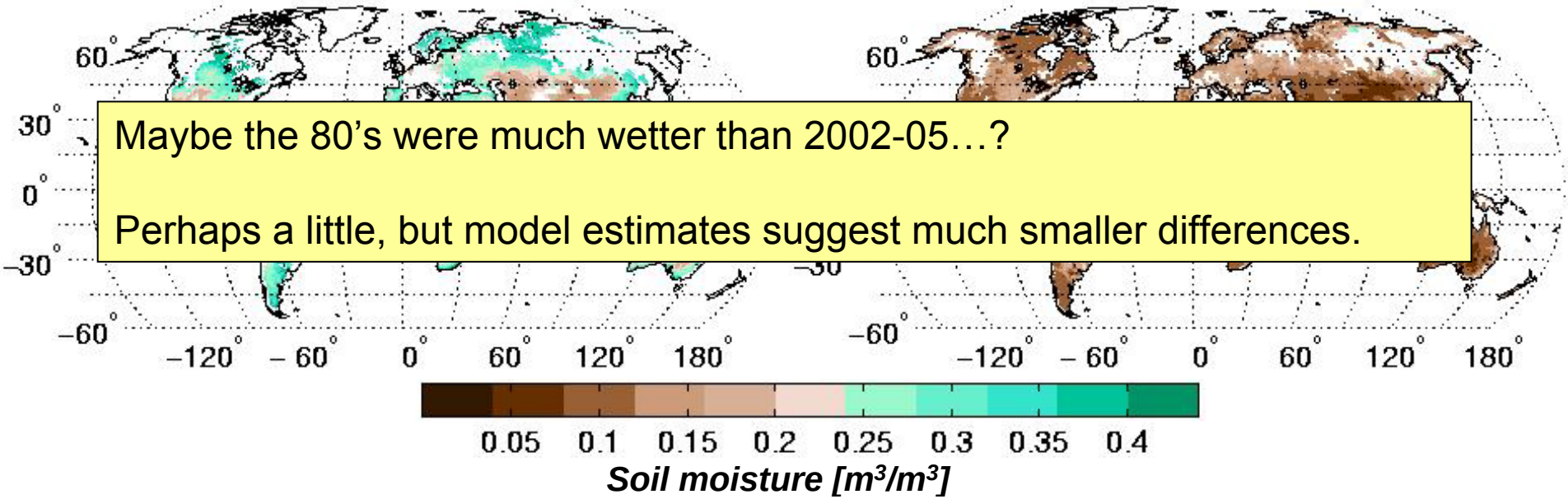
Soil moisture [m³/m³]

SMMR retrievals **much** wetter than AMSR-E retrievals.

Magnitude of differences comparable to dynamic range.

Satellite vs. satellite bias (time avg. soil moisture)

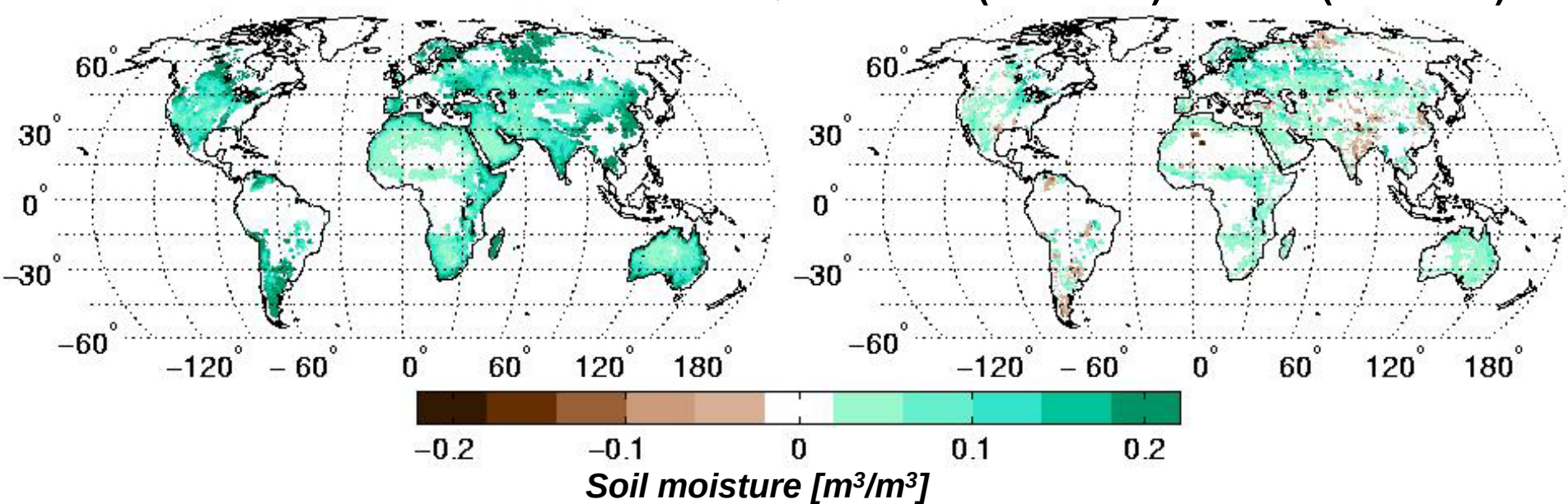
SMMR (1979-87) **AMSR-E (2002-05)**



Maybe the 80's were much wetter than 2002-05…?

Perhaps a little, but model estimates suggest much smaller differences.

SMMR minus AMSR-E **Model (1979-87) minus (2002-05)**

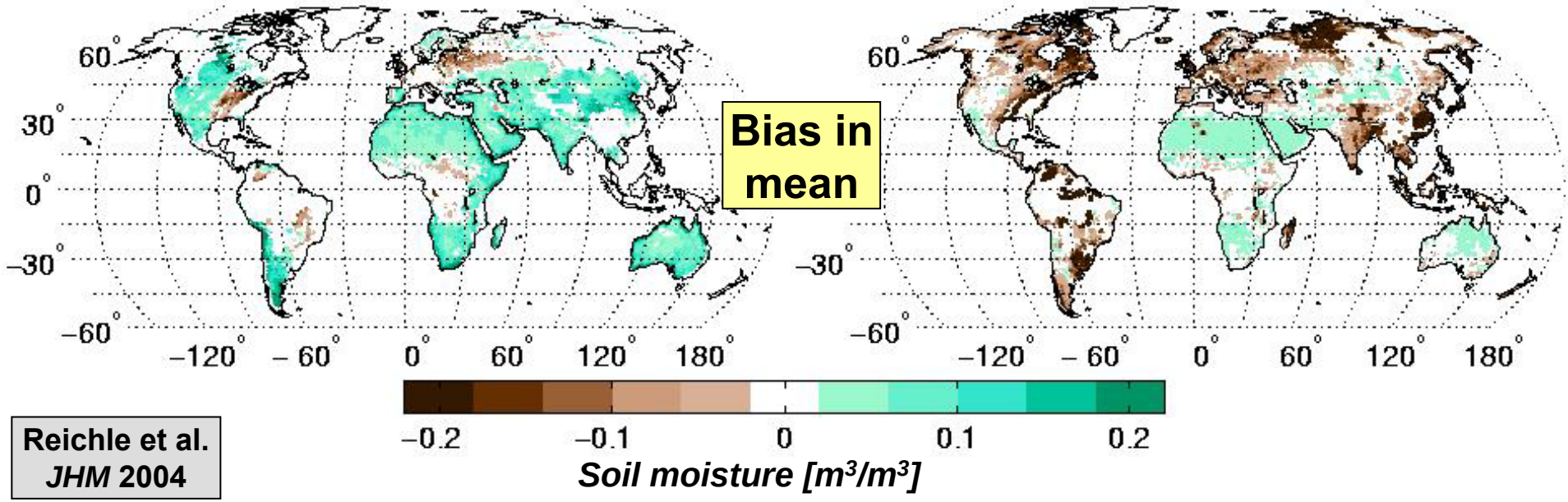


We found strong biases between AMSR-E and SMMR.

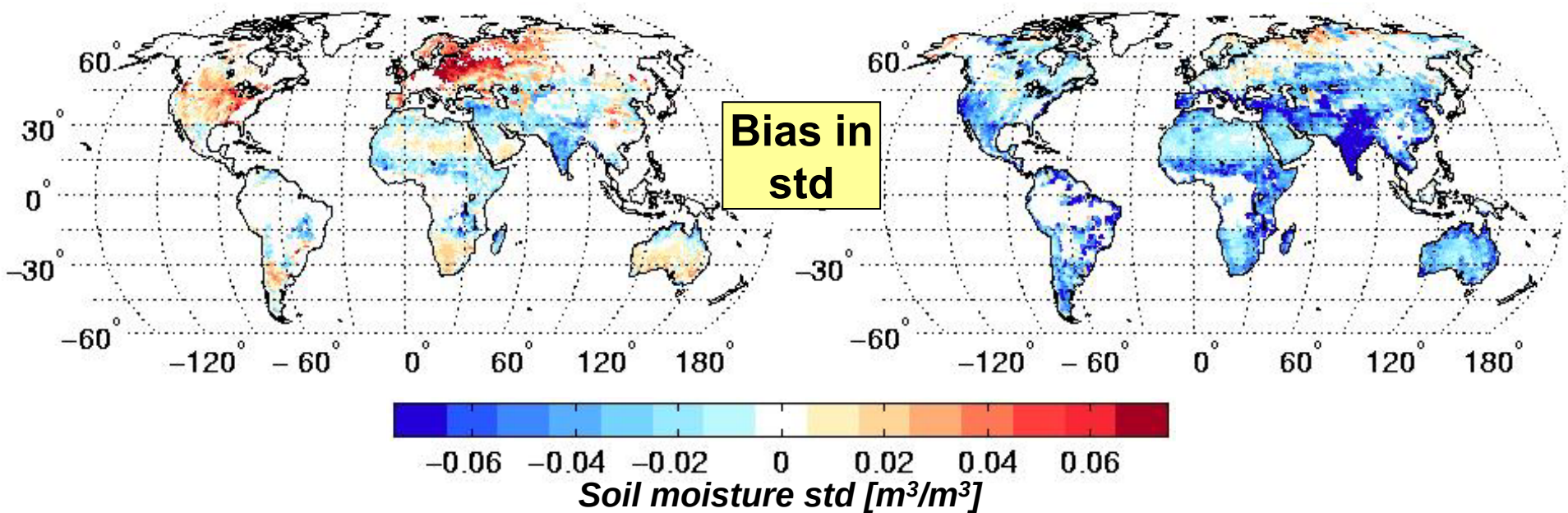
For assimilation, we are really interested in the **satellite vs. model** biases.

Satellite vs. model bias (soil moisture)

SMMR minus model (1979-87) **AMSR-E minus model (2002-05)**

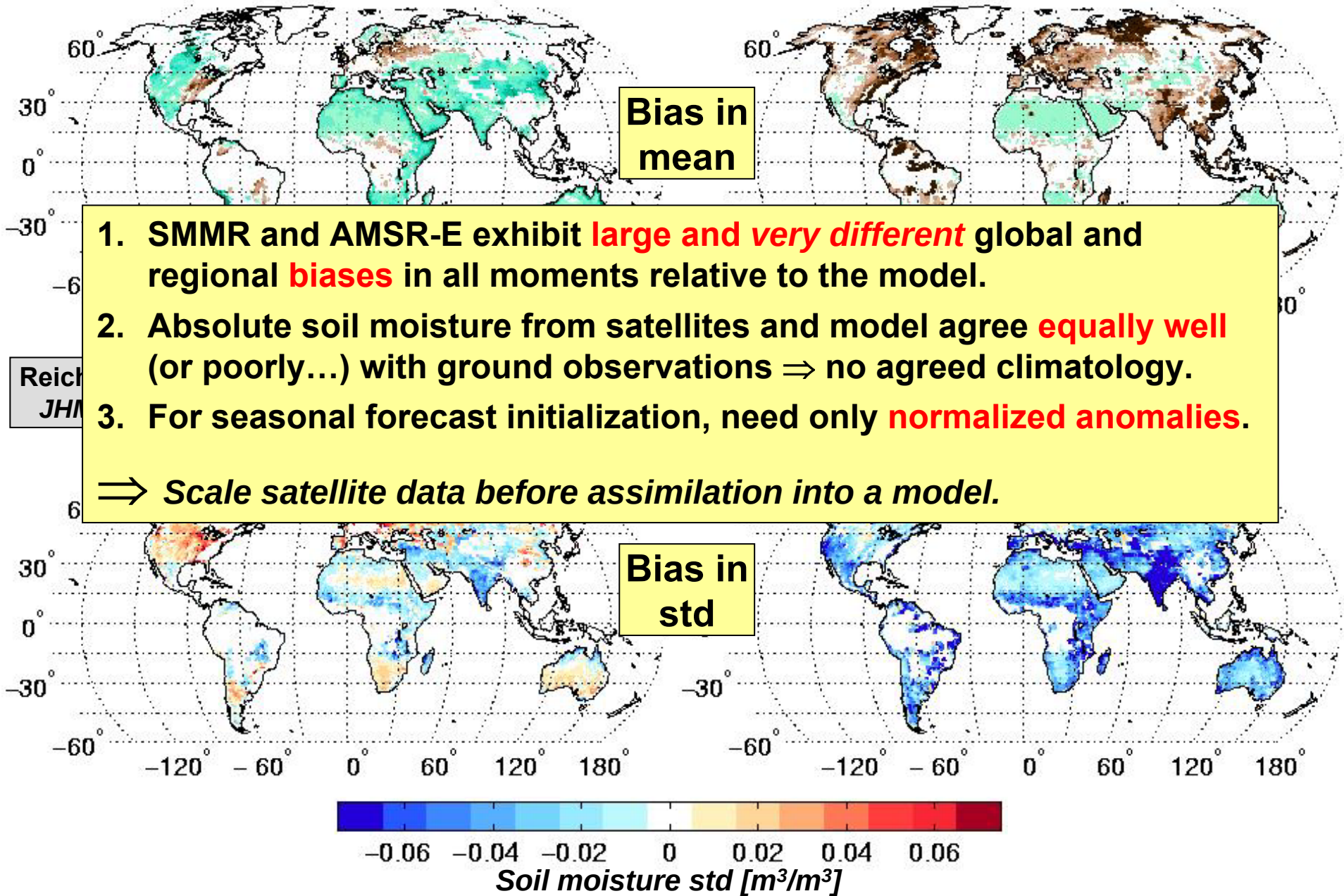


Reichle et al. *JHM* 2004



Satellite vs. model bias (soil moisture)

SMMR minus model (1979-87) **AMSR-E minus model (2002-05)**

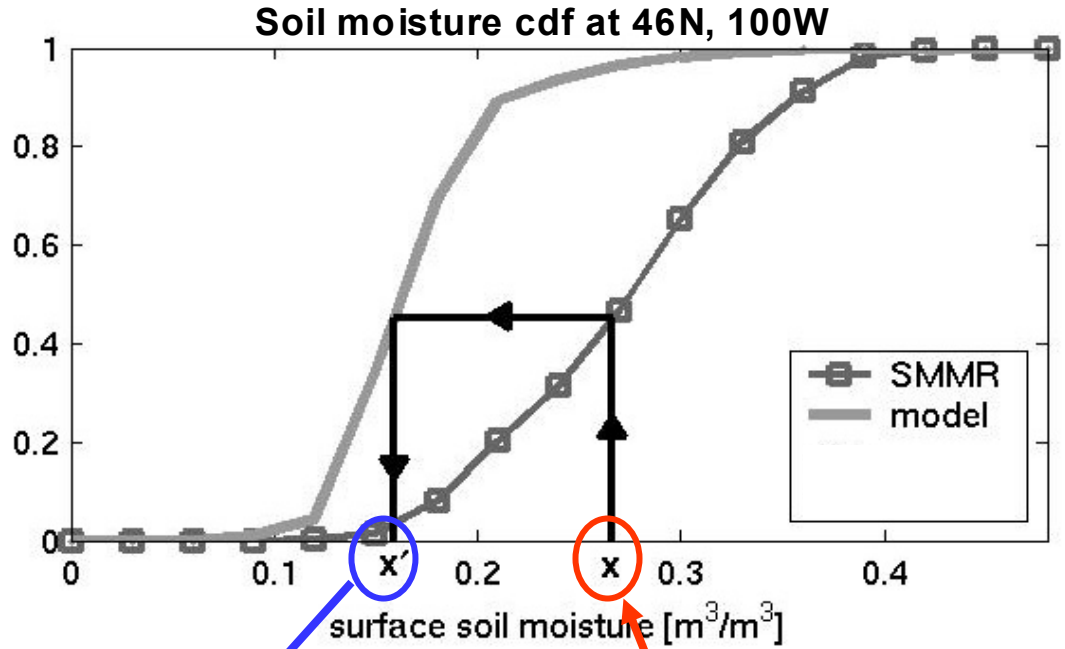


1. SMMR and AMSR-E exhibit large and *very different* global and regional biases in all moments relative to the model.
2. Absolute soil moisture from satellites and model agree equally well (or poorly…) with ground observations ⇒ no agreed climatology.
3. For seasonal forecast initialization, need only normalized anomalies.

⇒ *Scale satellite data before assimilation into a model.*

Reich
JHM

Soil moisture scaling for data assimilation

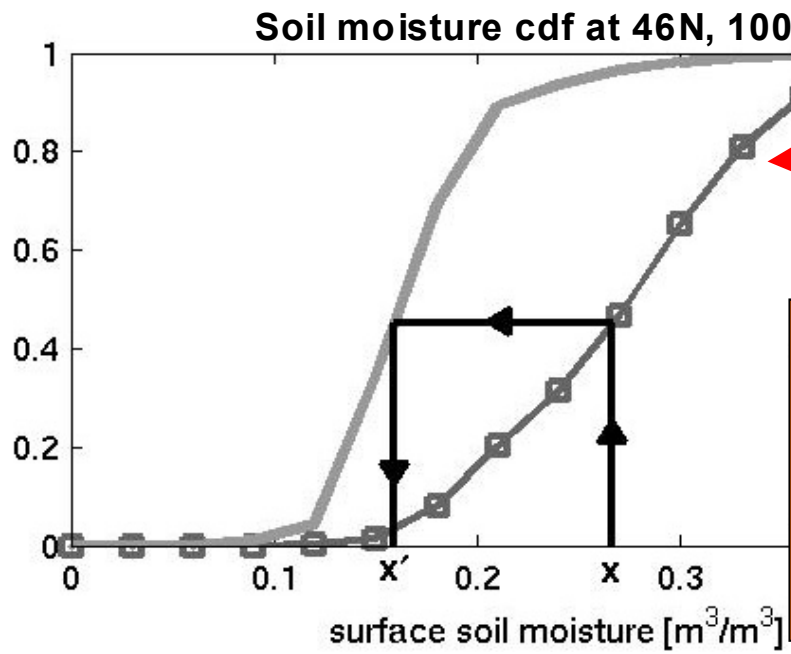


1. Find percentile of a given satellite measurement on the satellite's local climatological cumulative distribution function (CDF).

2. Find soil moisture that produces the same CDF value on the corresponding model CDF ⇒ "scaled" satellite measurement for assimilation.

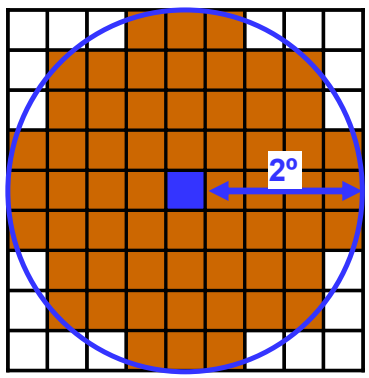
In short: Assimilate percentiles.

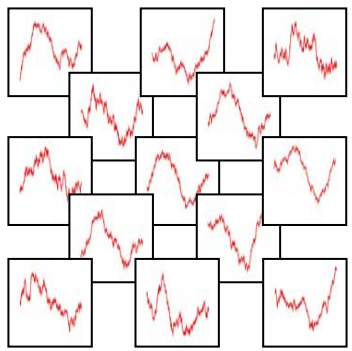
Soil moisture scaling for data assimilation



For "new" sensors:
- cannot use time series from historic satellites,
- long time series not immediately available!

<u>Solution:</u>
Approximate CDF from many 1-year time series at grid points within some distance from point of interest.

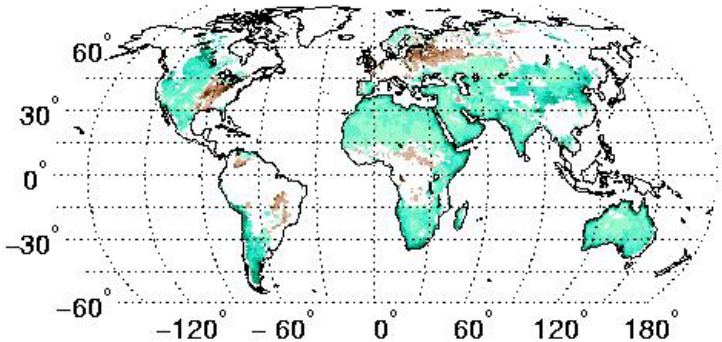




Soil moisture scaling for data assimilation (mean)

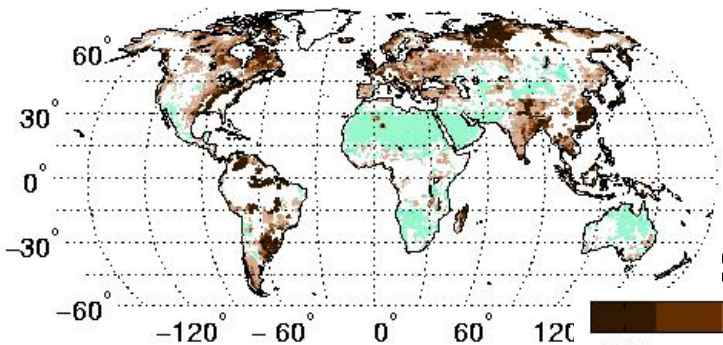
ORIGINAL multi-year data sets
(Satellite minus model)

SMMR

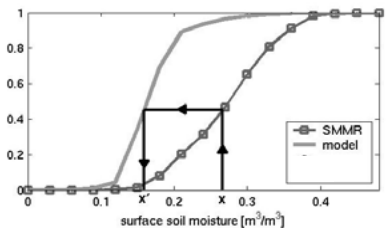


Reichle et al. *JHM* 2004

AMSR-E

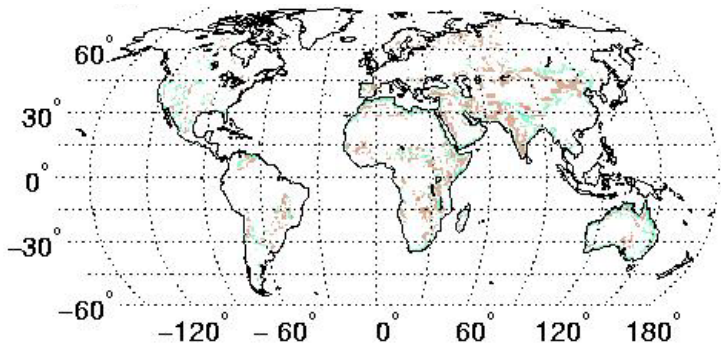


CDF scaling based on **1 year** of satellite data



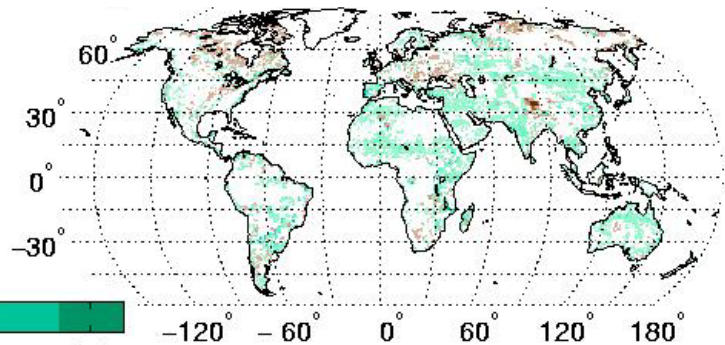
SCALED multi-year data sets
(Satellite minus model)

SMMR

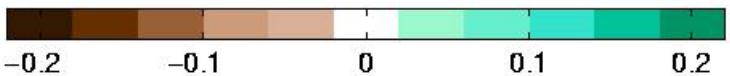


Reichle & Koster *GRL* 2004

AMSR-E



Soil moisture [m³/m³]

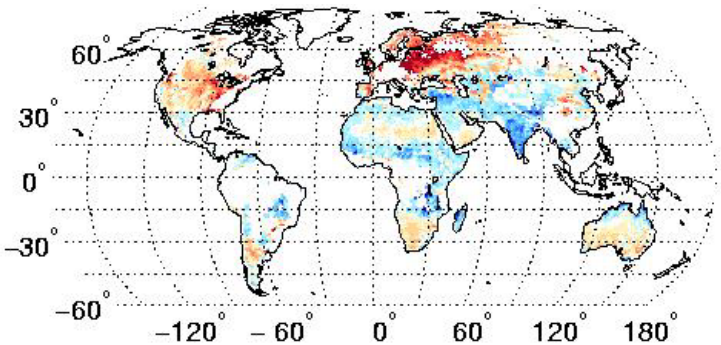


1 year of satellite data sufficient for considerable reduction in long-term bias.

Soil moisture scaling for data assimilation (std)

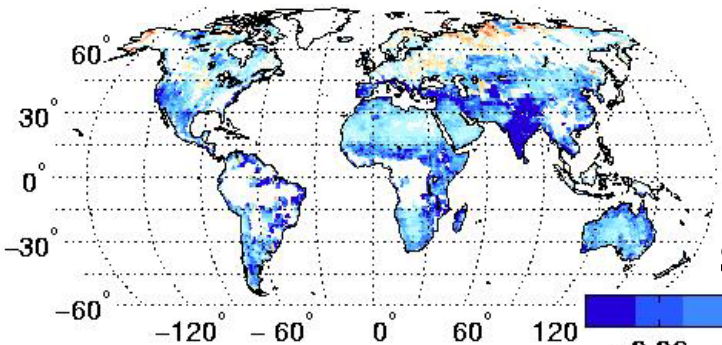
ORIGINAL multi-year data sets
(Satellite std minus model std)

SMMR

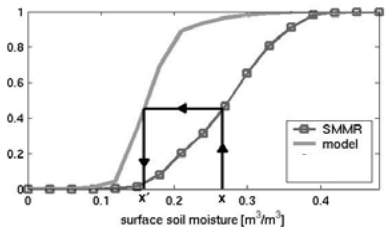


Reichle et al. *JHM* 2004

AMSR-E

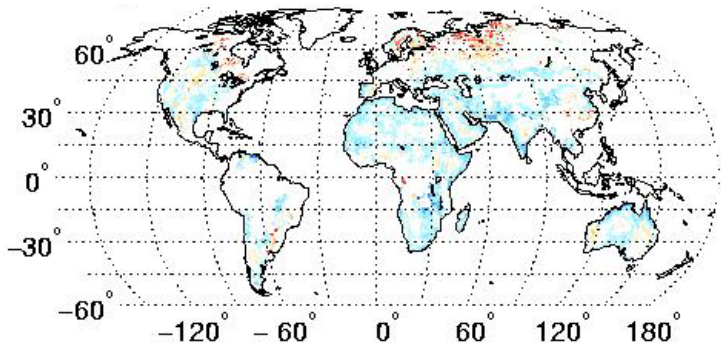


CDF scaling based on **1 year** of satellite data



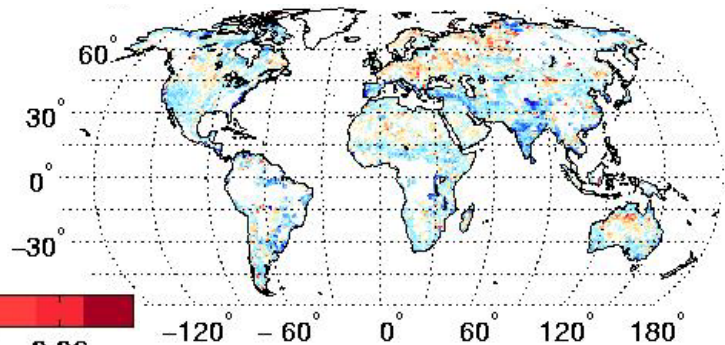
SCALED multi-year data sets
(Satellite std minus model std)

SMMR

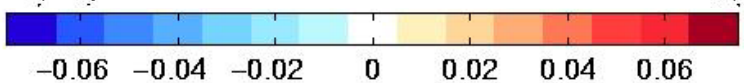


Reichle & Koster *GRL* 2004

AMSR-E



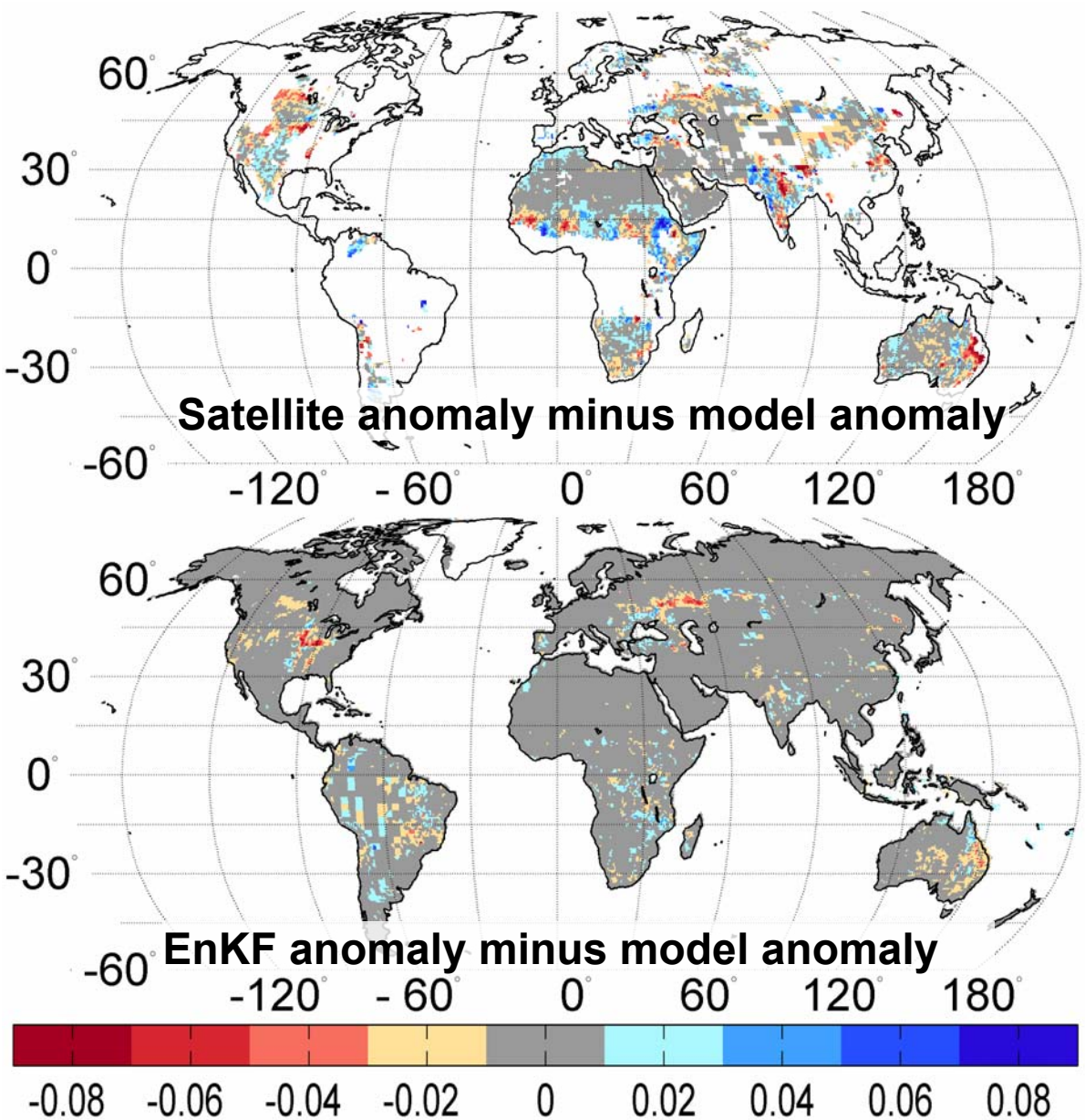
Soil moisture std [m³/m³]



1 year of satellite data sufficient for considerable reduction in long-term bias.

Assimilation of SMMR soil moisture

Impact of SMMR assimilation – July 1982

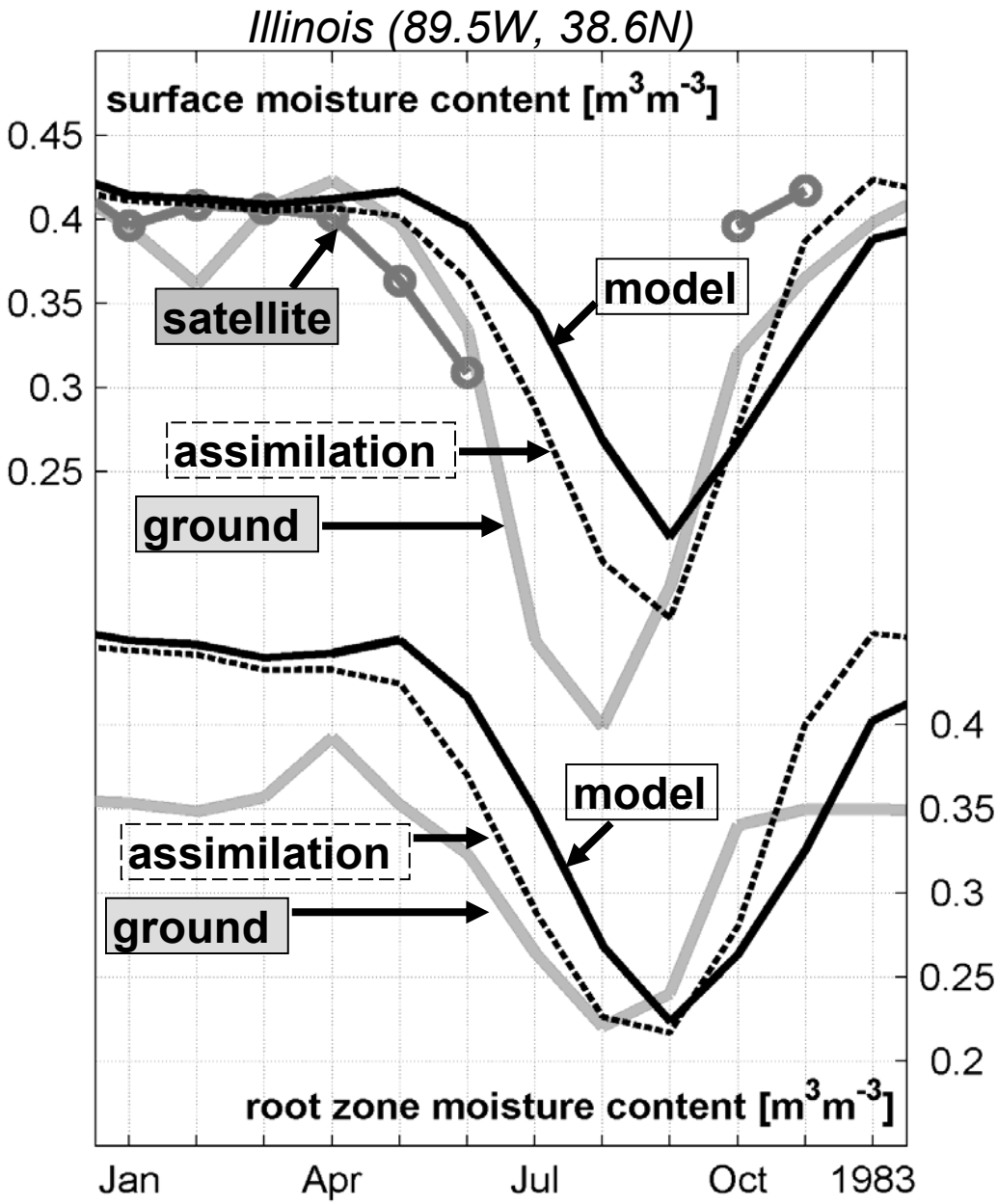


Assimilation product lies "between" SMMR and model.

There are interesting dynamical effects.

Reichle & Koster, GRL 2005

Validation against in situ data



Assimilation product has improved phase of annual cycle.

Reichle & Koster, GRL 2005

Validation of SMMR products against in situ data

		Time series correlation coeff. with in situ data [-] (with 95% confidence interval)			Confidence levels: **Improvement of assimilation over**	
	N	**SMMR**	**Model**	**Assim.**	**SMMR**	**Model**
Surface soil moisture	77	**.44±.03**	**.43±.03**	**.50±.03**	**99.7%**	**99.9%**
Surface anomalies	66	**.32±.03**	**.36±.03**	**.43±.03**	**99.9%**	**99.9%**
Root zone soil moisture	59	**n/a**	**.46±.03**	**.50±.03**	**n/a**	**97%**
Root zone anomalies	33	**n/a**	**.32±.05**	**.35±.05**	**n/a**	**80%**

Assimilation product agrees better with ground data than SMMR or model alone.

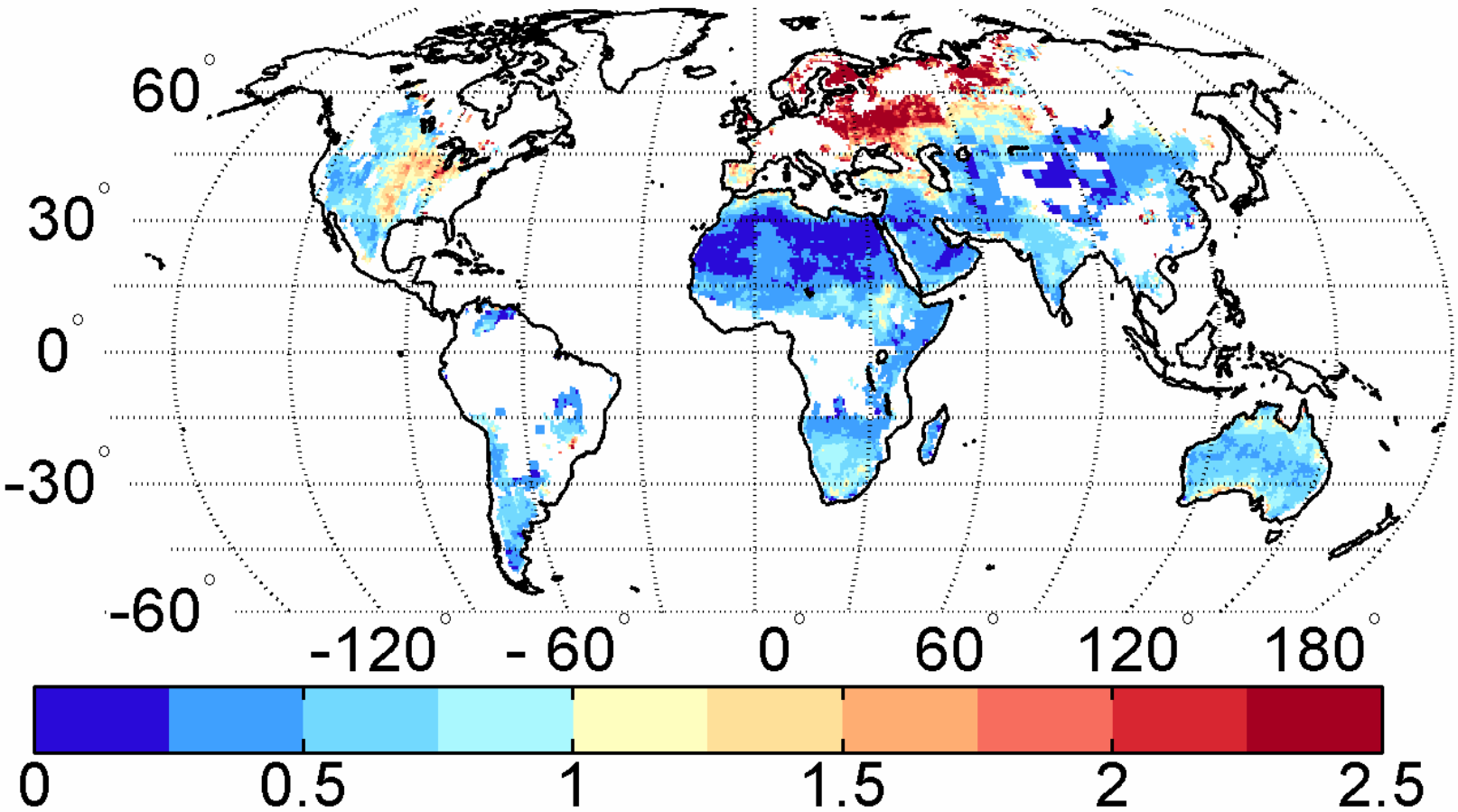
Modest increase may be close to maximum possible with *imperfect* in situ data.

Reichle & Koster *GRL* 2005

Assimilation of SMMR soil moisture

SMMR assimilation successful, but…

Variance of normalized innovations



Variance deficiency in dry climates, excess variance in wetter climates.

Potential for improvement by (adaptively) tuning model error parameters.

Reichle & Koster, GRL 2005

Validation of SMMR products against in situ data

	Time series correlation coeff. with in situ data [-]	Confidence levels: Improvement of
		lel
Surface moistur		**%**
Surface anoma		**%**
Root zo moistur		**%**
Root zo anoma		**%**

We are still working on a similar result for AMSR-E assimilation.

Success is perhaps more difficult to achieve because:

- we have only 3.5 years of AMSR-E (vs. 8.5 years of SMMR),

- AMSR-E retrievals based on X-band (vs. C-band for SMMR),

- GLDAS (2002-05) forcing data (vs. “Berg” 1979-87) might lead to improved (benchmark) model soil moisture.

Assimila alone.

Modest increase may be close to maximum possible with *imperfect* in situ data.

Reichle & Koster, *GRL* 2005

Conclusions (soil moisture)

Results:

Analyzed AMSR-E, SMMR, and model biases.

Soil moisture assimilation system must be capable of dealing with large biases.

SMMR assimilation improves land initialization.

Outlook:

Continue assessment of soil moisture estimates.

Further tuning and validation of AMSR-E assimilation.

Impact of SMMR and AMSR-E assimilation on seasonal predictions.

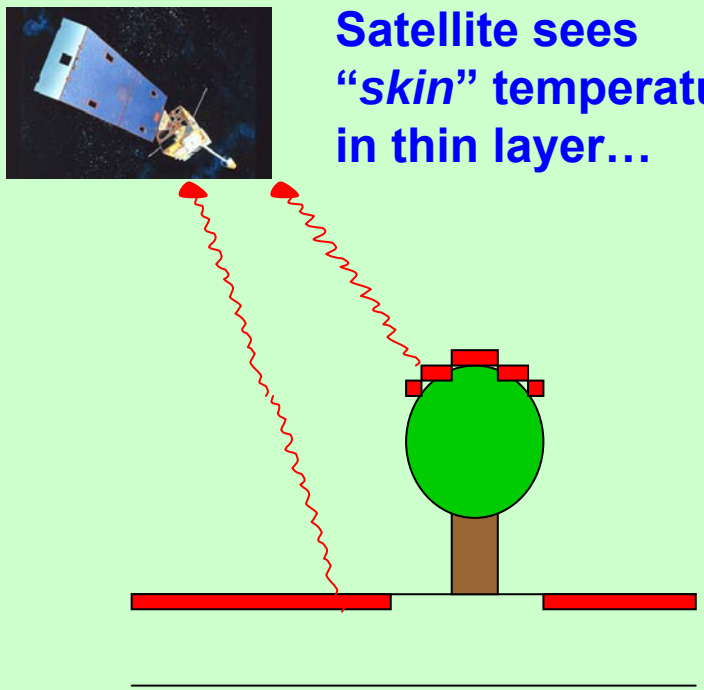
Outline

Motivation	Seasonal climate *predictability* associated with land surface conditions
Method	Land data assimilation with the Ensemble Kalman filter
Soil moisture	Satellite *observations*, *uncertainties* & data assimilation
Land surface temperature	Satellite *observations*, *uncertainties* & data assimilation
Conclusions	

Soil temperature assimilation

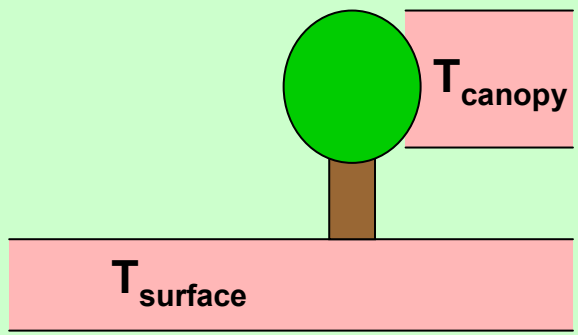
***Good news*: Abundance of satellite-based skin temperature from *infrared* and *microwave* sensors on *geostationary* and *polar-orbiting* platforms (NOAA-xx, MODIS, GOES, METEOSAT, GMS,…)**

***Problem 1*: Satellite *skin* temperature and model *surface* temperature may be inherently inconsistent (vertical).**



Satellite sees "*skin*" temperature in thin layer…

…whereas model "*surface*" temperature is typically a mixture of temperatures of thicker layers.

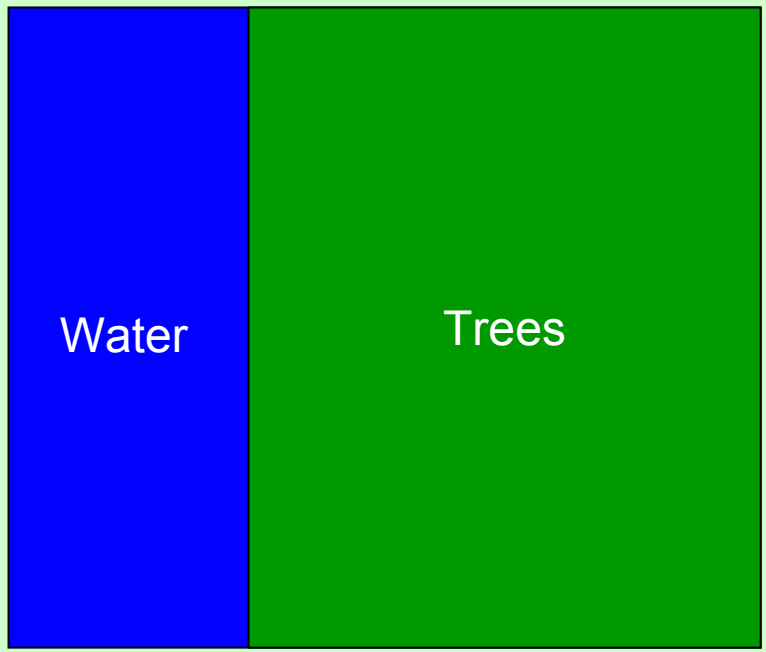


Soil temperature assimilation

***Problem 2*: Satellite *skin* temperature and model *surface* temperature may be inherently inconsistent (horizontal).**

Satellite "sees" a great variety of spatial heterogeneity…

…whereas a model is limited in the spatial variability it can represent.



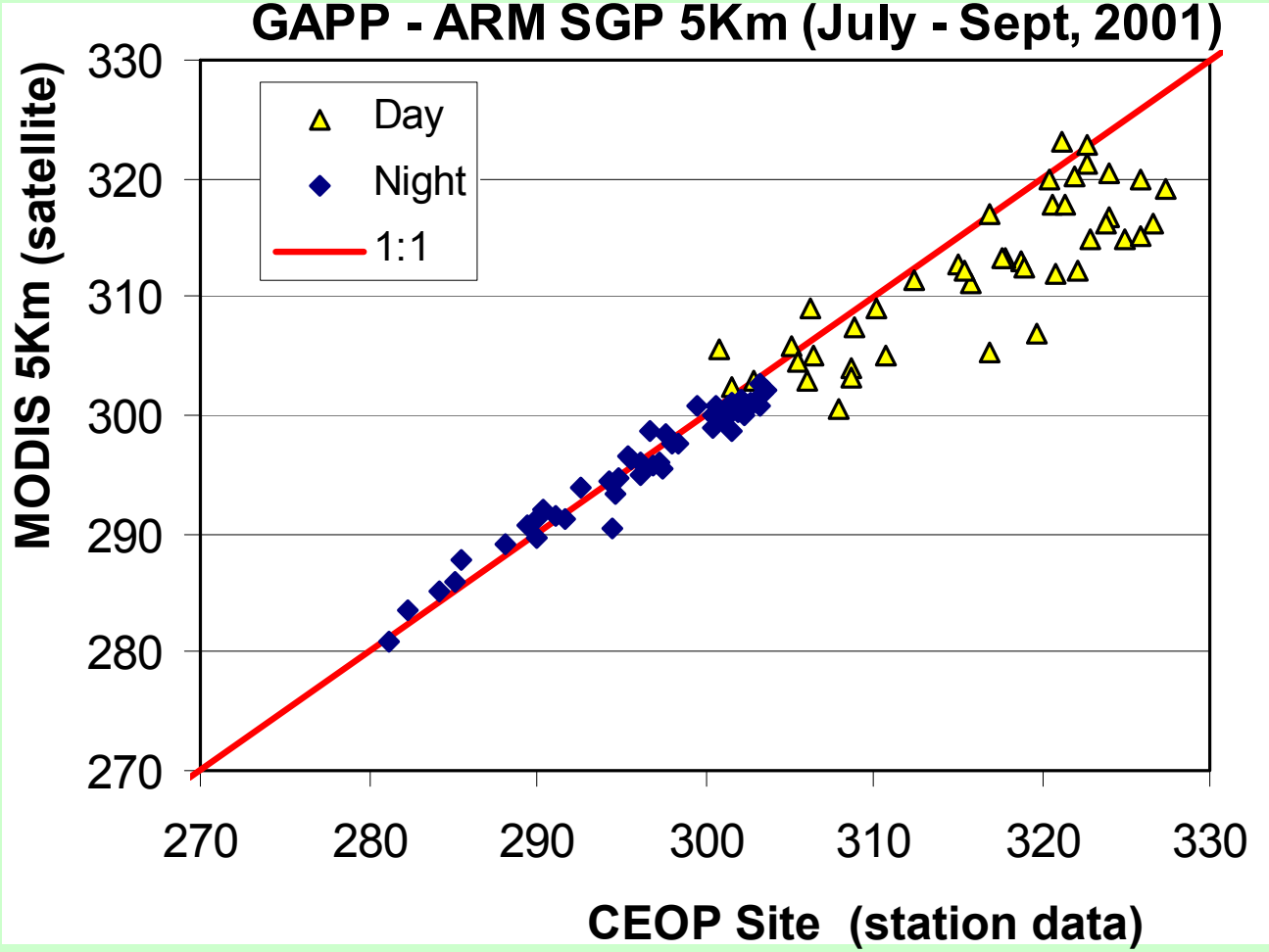
Shadows

Ice

Bare soil

Soil temperature assimilation

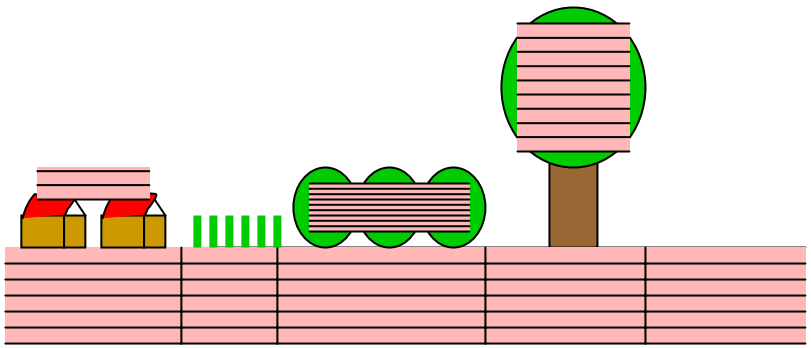
***Problem 3*: Satellite skin temperature may be sensor- or algorithm-specific.**



Three strategies for dealing with the inconsistency

1. Work on model and satellite algorithm until they are "perfect".

Problems include:



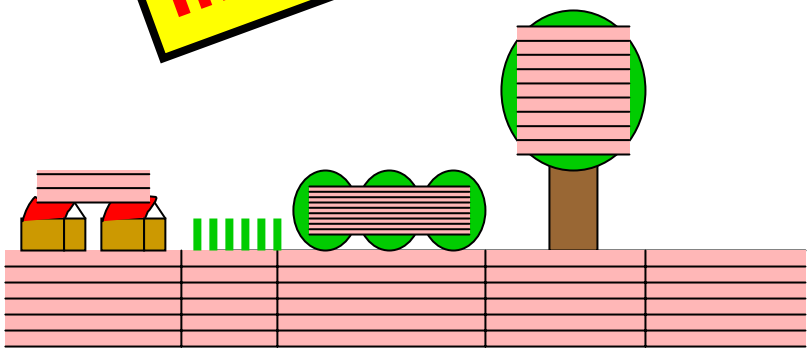
Thin layers create numerical stability issues.

Even with perfect model, lack complete spatial distribution of "correct" parameter values.

Three strategies for dealing with the inconsistency

1. Work on model tellite algorithm until they are "perfect".

In many ways, IMPRACTICAL!



Thin layers create numerical stability issues.

Even with perfect model, lack complete spatial distribution of "correct" parameter values.

Three strategies for dealing with the inconsistency

2. Ignore the inconsistencies and hope for the best.

Three strategies for dealing with the inconsistency

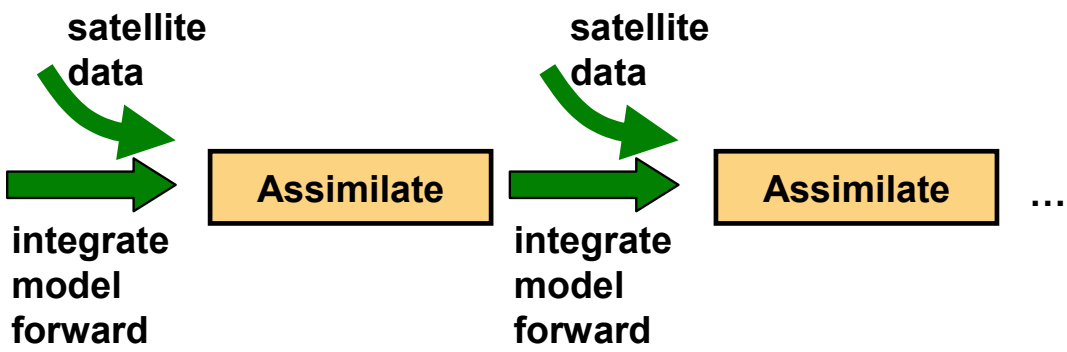
2. Ignore the inconsistencies and hope for the best.

Questionable results!
(as will be shown)

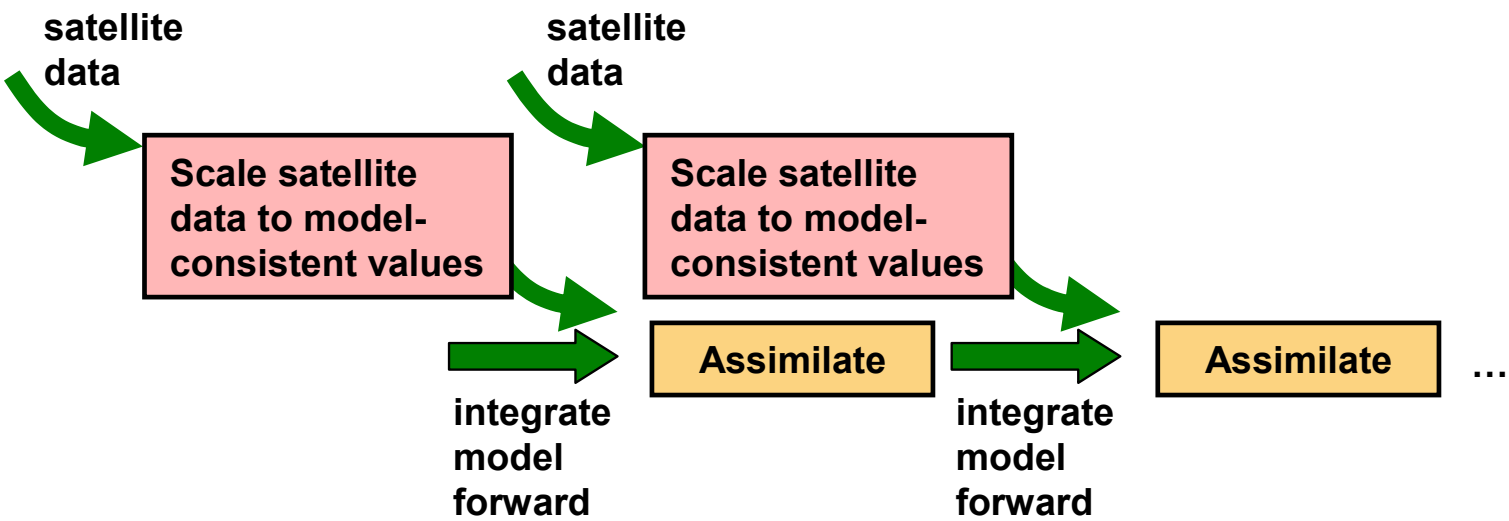
Three strategies for dealing with the inconsistency

3. Accept inconsistencies, scale the data prior to assimilation.

Instead of this…



…do this:



Three strategies for dealing with the inconsistency

3. Accept inconsistencies, scale the data prior to assimilation.

1) Get time series *mean μ* and *standard deviation σ* for satellite Tskin ("T_sat") and for model-based synthetic Tskin observations ("T_mod"), broken down by *diurnal cycle* and *month*.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
0z	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ
3z	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ
6z	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ
9z	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ
12z	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ
15z	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ
18z	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ
21z	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ	μ,σ

2) Scale satellite Tskin ("T_sat") into model climatology (std normal deviates):

$$\mathbf{T_sat_scaled = \sigma_mod/\sigma_sat \cdot (T_sat - \mu_sat) + \mu_mod}$$

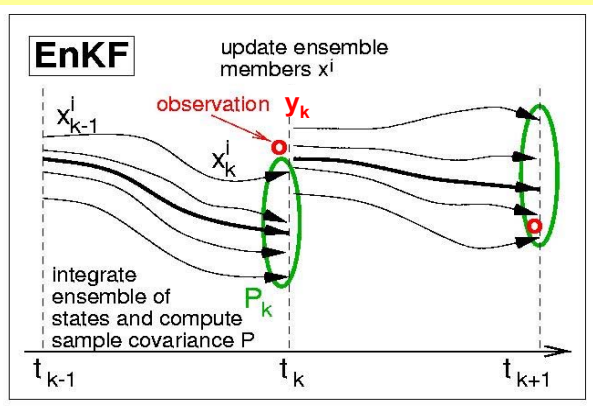
3) Assimilate scaled satellite Tskin ("T_sat_scaled").

Demonstration of scaling approach

Model:
NASA Catchment land surface model on 1 degree lat-lon grid.
(“off-line” – not coupled to atmospheric model)

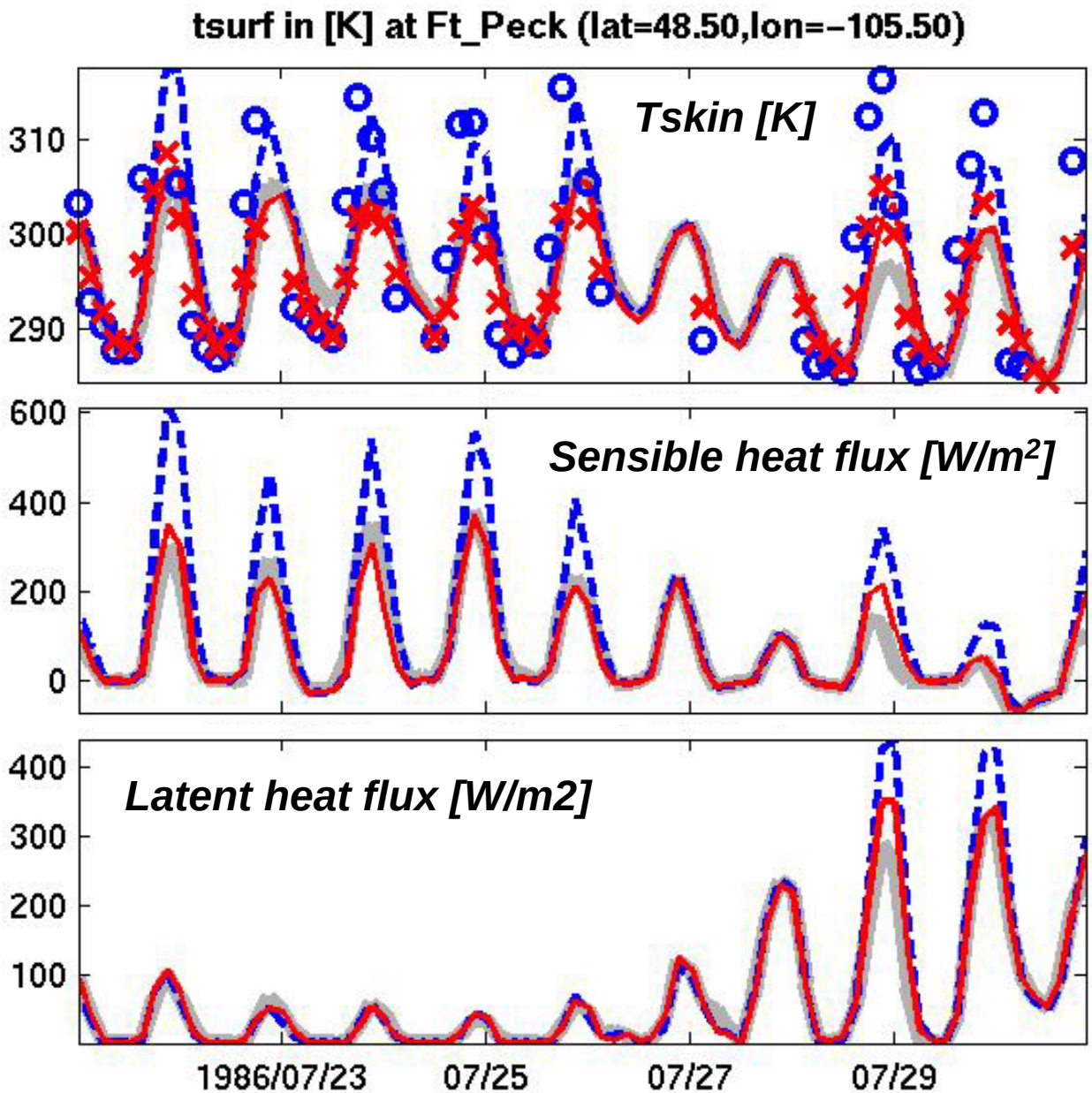
Surface meteorological forcing data:
- Global Soil Wetness Project (GSWP-2; 1986-95)

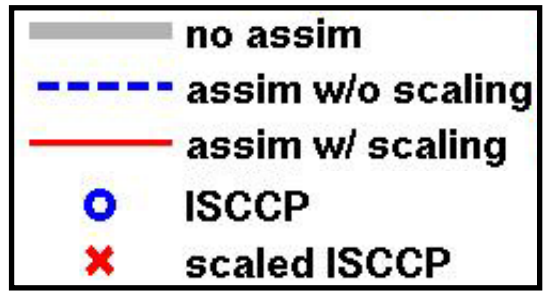
Satellite skin temperature:
- Int’l Satellite Cloud Climatology Project (ISCCP; 1983-2004)
  (NOAA-xx, GOES, METEOSAT, GMS,…)
- 3-hourly, mapped to 1 deg lat-lon grid
- clear-sky only!



Assimilation:
Ensemble Kalman Filter
(developed at NASA/GMAO)

A few DAYS in July 1986 at Ft Peck, MT, USA





Tskin mean and dynamic range from satellite and model differ. Assimilation w/o scaling increases peak Tskin.

When assimilating w/o scaling, model produces excessive sensible heat flux.

Latent heat flux also increases when soil moisture is available.

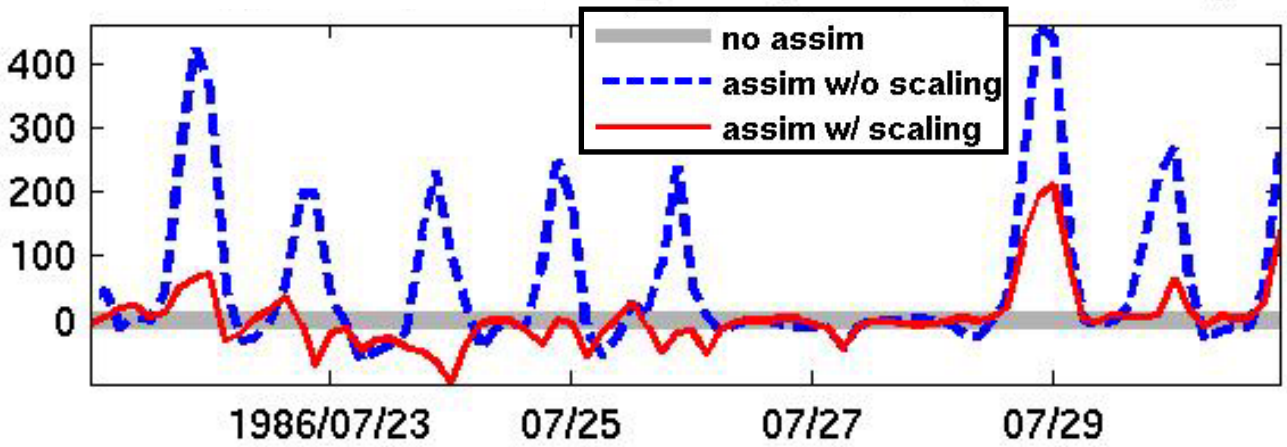
"Assimilation flux"

Surface energy balance:

G = Rnet - LE - H - "assimilation flux"

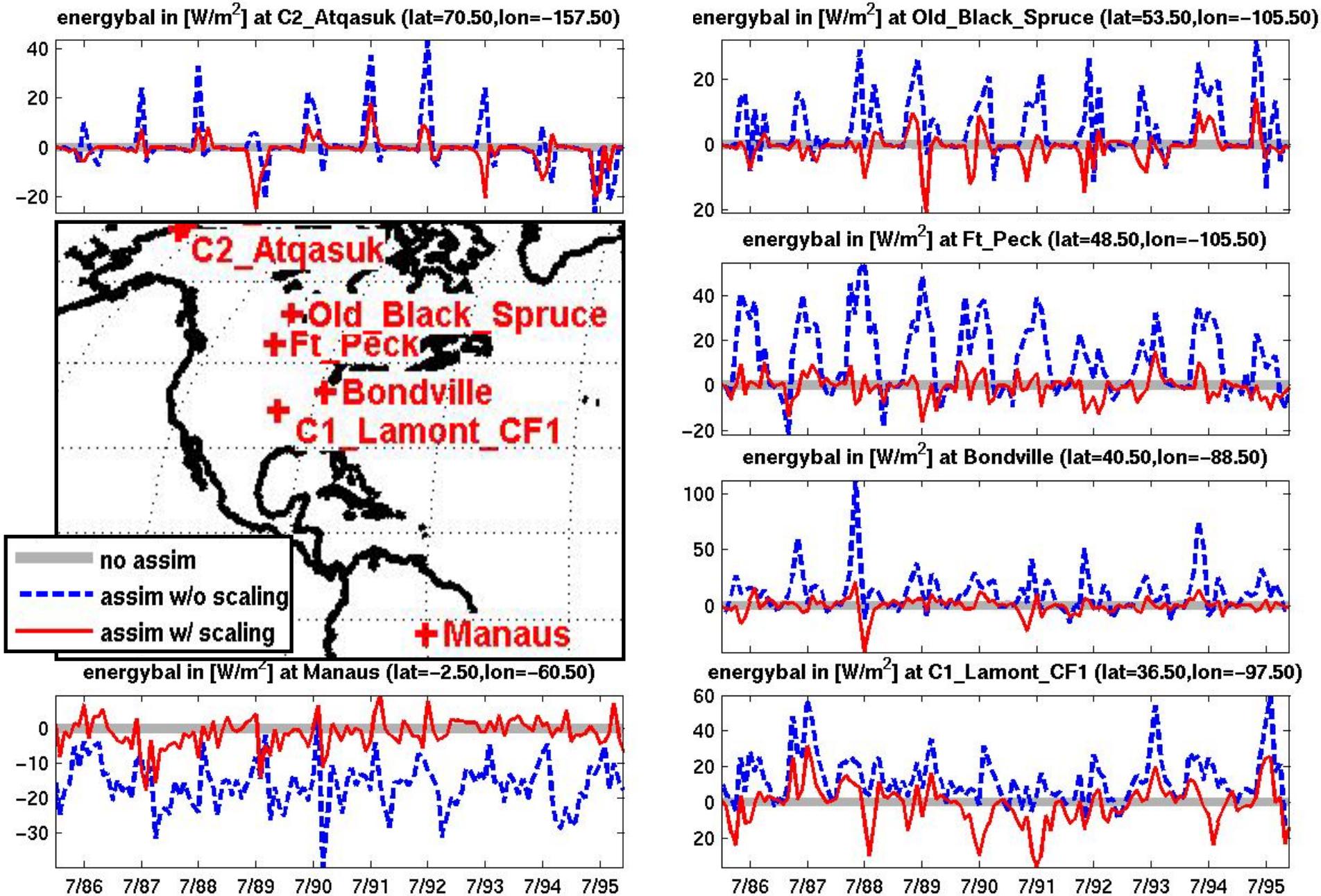
G = Ground heat flux *Rnet* = Net radiation *LE* = Latent heat flux *H* = Sensible heat flux	***Assimilation flux* =** **Added energy flux such that model pulls close to Tskin observations.** **Ideally small and white noise in time.**

***Assimilation flux [W/m²]* at Ft_Peck (lat=48.50,lon=−105.50)**

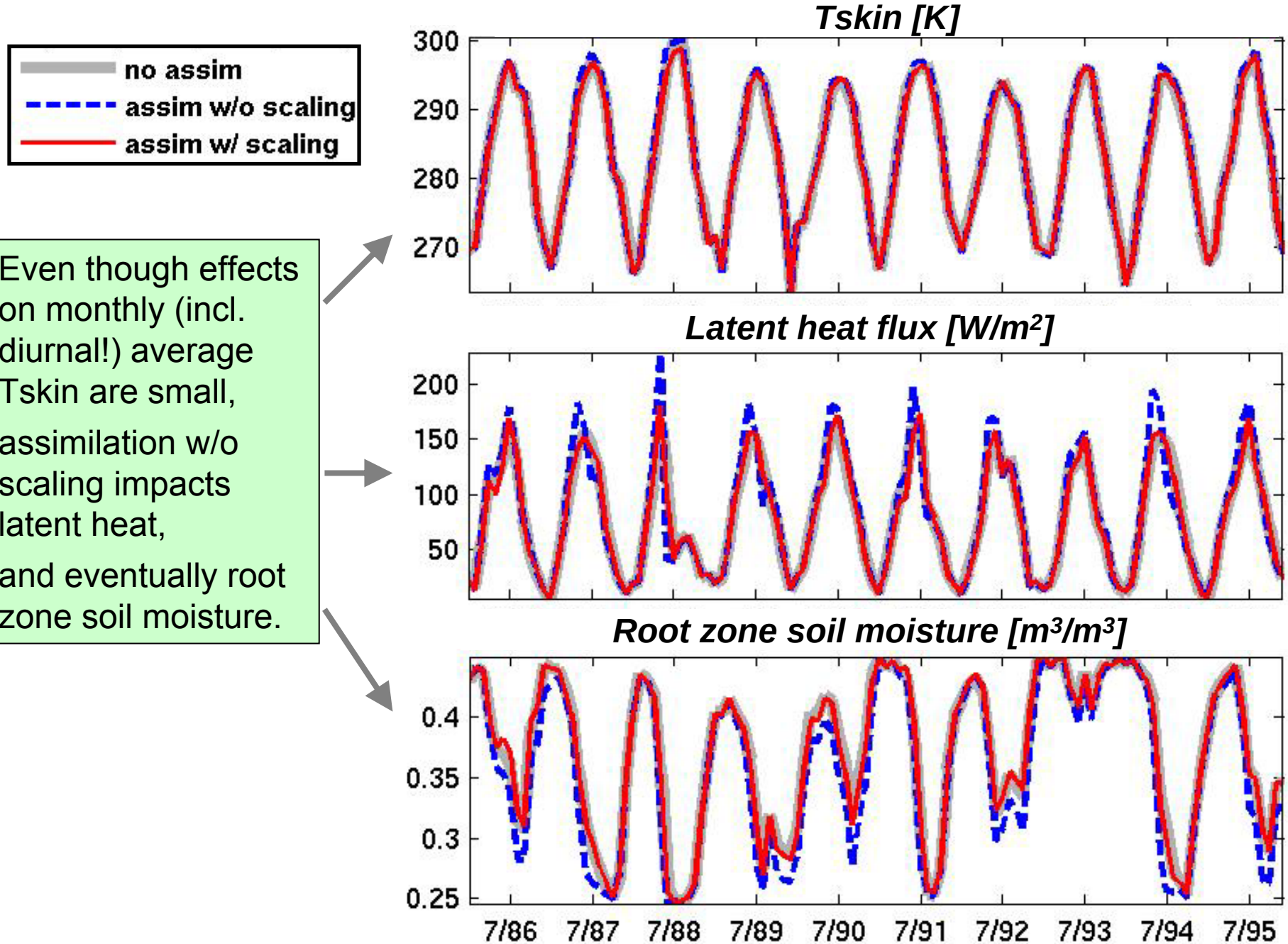


Excessive, non-white "assimilation flux" when assimilating w/o scaling.

MONTHLY average assimilation flux (1986-1995)



Impact on root zone moisture (Bondville, IL, USA, 1986-1995)



Even though effects on monthly (incl. diurnal!) average Tskin are small,

assimilation w/o scaling impacts latent heat,

and eventually root zone soil moisture.

Conclusions (land surface temperature)

1.) Differences between ***satellite skin*** and ***model surface temperatures*** are due to errors in, and inconsistencies between:

- satellite retrieval algorithm,
- model physics and parameterization,
- representation of spatial heterogeneity,
- vertical resolution, …

2.) ***Assimilation of Tskin retrievals must consider differences between satellite and model climatologies.***
Otherwise, excessive assimilation fluxes are required to force the model Tskin towards the retrievals, and ***excessive and unrealistic sensible and latent heat fluxes are generated***.

3.) ***Avoid such negative effects by scaling the satellite Tskin into the model climatology prior to assimilation.***
Because of the strong diurnal and seasonal cycles of Tskin, scaling statistics must be broken down by month and hour of day.

Conclusions

Land surface conditions may contribute to predictability at sub-seasonal time scales.

EnKF is suitable for estimating land surface conditions.

Large biases between satellite and model estimates can be addressed with a scaling approach ("assimilate percentiles").

Assimilation of SMMR soil moisture retrievals provides superior soil moisture estimates when compared to the model or the satellite alone.

Much is left to be done…
… demonstrate skill with AMSR-E & Tskin assimilation,
… develop snow assimilation,
… assimilation in the coupled land-atmosphere system,
…

The End